Piriyapongpipat *et al.*
Smart Learning Environments (2024) 11:20
https://doi.org/10.1186/s40561-024-00307-8

Smart Learning Environments

RESEARCH

Open Access

An alternative approach to ontology-based curriculum development in higher education

Pattamaporn Piriyapongpipat[1], Sally Goldin[2] and Nadh Ditcharoen[1*]

*Correspondence:
nadh.d@ubu.ac.th

[1] Faculty of Science, Ubon Ratchathani University, 85 Sathonlamark Road, Warin Chamrap District, Ubon Ratchathani 34000, Thailand
[2] Electrical and Computer Engineering Department, CMKL University, 1 Soi Chalongkrung 1, Ladkrabang, Bangkok 10520, Thailand

Abstract

Global trends in higher education emphasize the development of curricula that offer greater responsiveness to learners. Creating flexible and responsive curricula will require additional support systems for curriculum management. The first step toward sustainably developing this kind of system is to represent essential curricular information in a way that allows sharing common components across various work processes within the educational environment. The current research implements a new approach for representing curriculum components, by systematically analyzing external source data to extract the basic knowledge, skills and dependencies which then become objects into an ontology. The resulting ontology should act as a computationally-accessible model of the curriculum with sufficient information and usable quality. This paper describes a trial implementation of our approach using actual curriculum documents. Results from performance metrics and expert evaluation validate the proposed strategy and suggest that the approach is feasible for real-world practice.

Keywords: Ontology, Knowledge modelling, Curriculum development, Higher education, TQF

Introduction

At present, there are numerous ongoing efforts to change higher education in order to provide the best benefit and opportunity for learners. Government policies, pedagogical research and practical educational trends all indicate a shift from traditional, formal, course-based education to flexible, individualized, learner-centered activities. There is a growing consensus that higher education should focus on building the necessary competencies to support lifelong learning (Bakhshi et al., 2017; Hicks, 2018; The Government of Thailand, 2020; O'Malley & Warden, 2022; UNESCO, 2022). One of the keys to success in this transition is curriculum design. The design must accommodate the diversity of domain knowledge needed to adapt to rapid changes in technology, business and society. Curriculum management needs to be more flexible and responsive to support these new directions and must be appropriate to new education ecosystems, such as national credit banks, micro-credentials, nano-degree programs and online universities (Andrade, 2018; Human-Hendricks & Meier, 2020; Heggart, 2022; World Economic Forum, 2022; Vreuls et al., 2022). In practice, however, curriculum management still struggles to effectively deliver the output as desired, and this continues to be a significant barrier to the development of higher education. The main obstacle is the complex and time-consuming process of modifying curriculum structure. To create or to revise a curriculum requires detailed information on both the curriculum content and the learning process. Even modest revisions may have to follow complicated protocols due to institutional regulations. In addition, most of the curriculum development workflow still relies on manual operations with an unfortunate lack of automated support infrastructure (Costandius & Bitzer, 2015; Chakraborty et al., 2016; Ávila et al., 2017; Barrier et al., 2019; Tractenberg et al., 2020; OECD, 2020).

Accordingly, the idea of using computer systems to assist in curriculum management has been widely discussed as a partial solution (Cheng & Nunes, 2022). The ontological approach is one proposed method which has shown promise as a foundation for a curriculum management support system. An ontology is a type of structured data which can capture complicated relationships among information elements. Ontologies support transformations and inference processes involving the knowledge they represent. The flexibility, extensibility and automated processing provided by ontologies fit well with the likely requirements of future curriculum development platforms (Chi, 2009; Lu & Zhang, 2011; Ferna Lndez-Breis et al., 2012; Komenda et al., 2015; Dehne & Kiy, 2019).

Our interest is on a complete integration of ontology into curriculum management in order to contribute to solutions for advanced curriculum design (Burtscher et al., 2015; Lightfoot, 2006; Mark, 2022). However, attempts to utilize ontologies for curriculum representation have been hampered by labor and time required to transform curriculum information into structured digital data. Typically, research on ontology-based curriculum management has involved two processes. The first process is to transform curriculum into machine-readable form and the second is to incorporate the data into a curriculum design application. Most of the prior studies have focused on the second process. The first process is rarely a priority. Previous researches usually created an entirely new set of data structures to fit with the design of the curriculum management system then manually prepared or arranged input data in the database (Chung & Jeongmin, 2016; Bussemaker et al., 2017; Nuntawong et al., 2017; Zouri & Ferworn, 2021). This approach is not only labor-intensive, but also makes it hard to share results across systems or to extend systems created by other groups, since the schema for each data representation is customized and unique.

In order to take the first step toward our main focus, a suitable methodology for developing curriculum as an ontology is required. Curriculum data which has been transformed into an ontology can be called a "curriculum ontology" (Liu et al., 2014). To provide support for systems that facilitate the expansion of new education ecosystems, the curriculum ontology should have a generic schema that allows it to represent the knowledge and structure of many different curricula. It must provide common components which can drive all the tasks involved curriculum management processes, as well as contributing to the systematic acquisition of new knowledge components. The schema should also be specific and concrete enough to support the creation of computer-based tools for curriculum ontology generation. The present research aims to fulfill these requirements. Our study presents a generic schema and demonstrates a usable approach for analyzing general curriculum content by considering essential curriculum elements as objects, then mapping the objects with the ontology components. This allows us to create a curriculum ontology based on systematic, reproducible processes using the generic schema. We demonstrate the practical utility of our schema by implementing a system to automatically derive curriculum ontology from general curriculum data.

Our expected outcome is an ontology which can practically and comprehensively represent the semantic content of the curriculum. Given this implementation, we can ask: does the outcome as a curriculum ontology provide sufficient quality to represent actual curriculum content in term of fundamental curriculum components?

In "Methodology" section, we present summary of the knowledge from previous works and our methodologies employed to address this research question. The results obtained from our methodologies are presented and discussed in "Results and discussion" section. In "Conclusion and future work" section, we give our final conclusion and outline our future work.

Methodology

To analyze the fundamental components of a curriculum, "Curriculum structure model" section presents the concept of fundamental components capable of representing a curriculum in a comprehensive manner. "Curriculum ontology" section presents the adapted method for transforming the concept discussed in the first section into ontology components and creating the ontology-based curriculum. The implementation procedures involved converting curriculum data from curricular documents into curriculum ontology are presented in "Implementation" section. The Thailand Qualifications Framework (TQF) documents are considered as the curricular documents utilized in this study. In "Evaluation" section presents the procedures for assessing the curriculum ontology derived from the implementation.

Curriculum structure model

In general, a curriculum consists of different information categories. Common categories which are central to almost every curriculum are the learning objectives and the subject matter. According the most common view, a curriculum defines a series of courses as a formal academic plan prepared for the students. Completion of the plan leads to achievement of the program's final goals (e.g. a degree) (Sand et al., 1960; Kelly, 2004; Oliver et al., 2008). Each course in the curriculum presents a set of objectives which describe what the students must accomplish in order to complete the course. A course objective typically consists of some learning object nouns and some learning behavior verbs to indicate the intended outcome of learning (Ferguson, 1998; Kennedy et al., 2009). The noun is a knowledge topic in the lesson, while the learning behavior verb is a level of cognition to show the action of achievement, for instance "understand", "analyze" or "apply". The noun and verb together define a learning objective. Hence, at the most basic level, every curriculum can be viewed as a hierarchical structure of courses, course objectives, learning topics and levels of cognition. In our proposed model, this structure is viewed as layers of data: the courses layer, the objective layer, and the knowledge layer.

Figure 1 illustrates the three layers of the curriculum structure in expanded perspective. The objects in the courses layer serve as frames or containers that define the boundaries for objects in the knowledge and learning objectives layer. The topic objects and

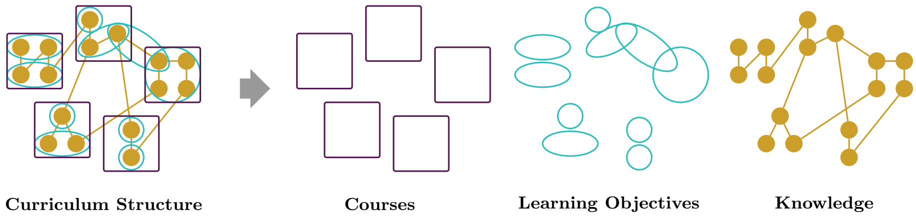

Fig. 1 Curriculum structure: expanded perspective

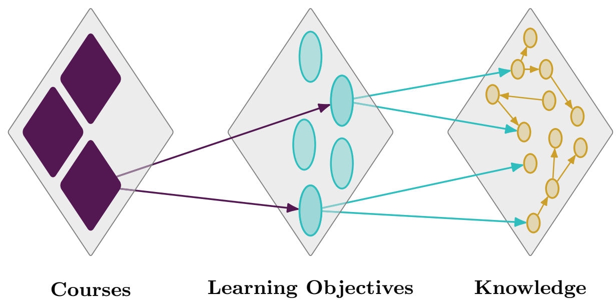

Fig. 2 Curriculum structure: in-depth perspective

their relationships inside the knowledge layer convey essential data that represents the subject matter and the sequence of study. The objects in the learning objectives layer depict the learning objectives which comprised of the learning topics and the levels of cognition. Figure 2 illustrates the curriculum structure in depth displaying the relationships between layers.

These three layers offer a systematic and well-specified object in different schema for distinguishing the basic components of a curriculum. The four basic components we have defined are *Course*, *Learning Objective*, *Cognitive Level* and *Topic*, called *Curriculum Objects*. We call semantic model which is drawn from these components the *Curriculum Structure Model* (Peckham & Maryanski, 1988; Robinson et al., 2015). Figure 3 shows relationships between different aspects of a curriculum and these curriculum objects.

The knowledge layer provides a path of the high-granularity learning topics to represent the learning sequence defined for the curriculum. The topic objects in the layer are ordered, such that the earlier units often involve simple principles which lay the foundation for the more complicated material found in later units (Doignon, 2014). For example, in a data structures course in computer science, Arrays and Linked Lists topics are normally studied first followed by the tougher topics of Stacks and Queues which in turn are followed by the even more sophisticated Trees and Graphs topics. The arrows in Fig. 3 within the knowledge layer indicate the prior entities in the curriculum-specified learning sequence. The Stacks and Queues topic continues from the Linked Lists topic, meaning that the Linked Lists lesson is taught before the Stacks and Queues lessons.

The learning objectives layer expresses the relationships between the learning topics and the cognition levels which explicitly appear in the curriculum. For example, a statement of one course objective says "To understand basic concepts of Stacks and Queues". The information associated with this objective will include the "to understand" relation,

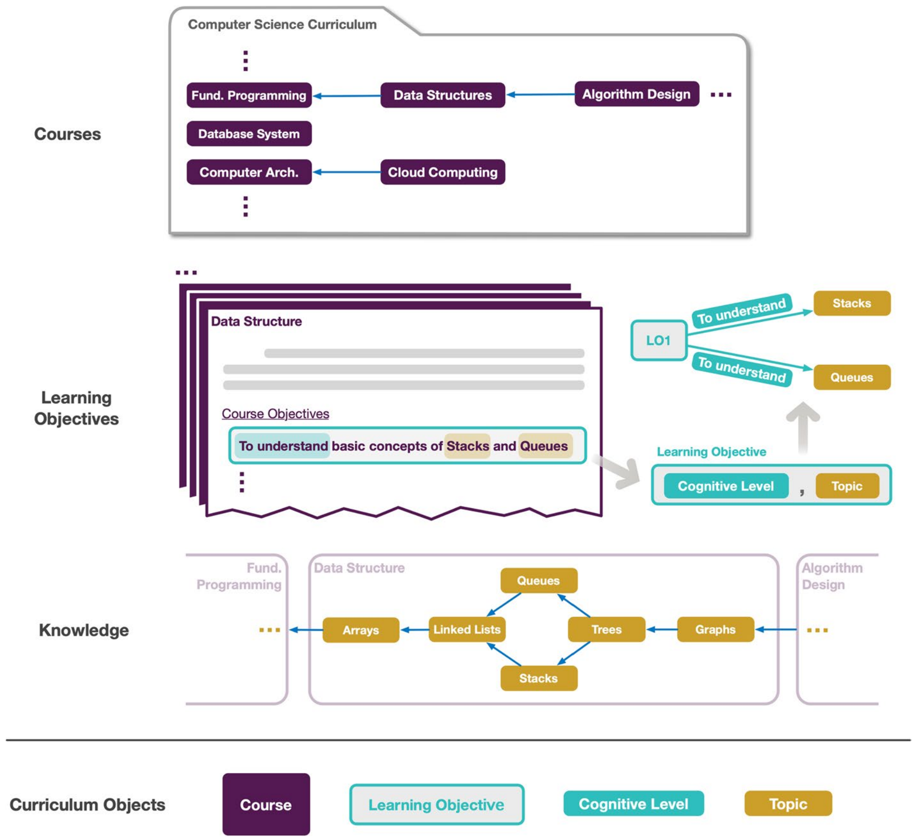

Fig. 3 Example of curriculum structure model interpretation

addressed to the Stacks and Queues topics. In Fig. 3 LO1 comprises two relationships, which are modeled as two classes in our ontology: the "to understand" relation to the "Stacks" topic and the "to understand" relation to the "Queue" topic.

The objects in the courses layer represent the courses offered in the curriculum. The prerequisite conditions indicate the continuation among courses in the curriculum. For example, in Fig. 3, the Algorithm course continues from the Data Structure course, meaning that the Data Structure course is a prerequisite of the Algorithm course for this curriculum. The curriculum structure model demonstrates these continuities in the knowledge layer, as depicted in Figs. 1 and 3. The aim is to organize the study sequence on a single layer, reflecting the order of learning topics between courses and facilitating backward enquiry.

Curriculum ontology

An ontology is a formal representation using concepts, domains and their relationships. Their elements capture knowledge about a semantic domain in a generally usable form (Noy & Mcguinness, 2001). To construct an ontology from curriculum data, we must to capture a set of triples, which represents knowledge of an association between two curriculum elements (element-relation-element), as the fundamental unit of the ontology (Atapattu et al., 2017; Fiallos & Ochoa, 2019). The mutual

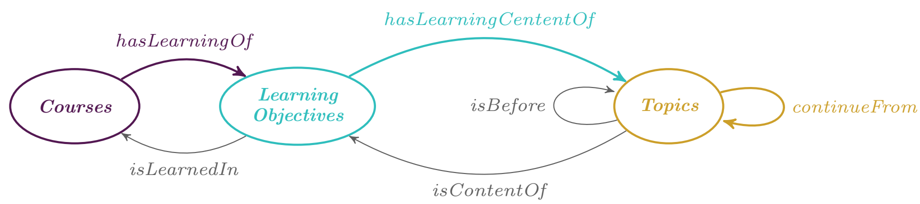

Fig. 4 High level classes and object properties of curriculum ontology

Table 1 Object property description of curriculum ontology

Object property	Domain	Range	Inverse Property
continueFrom	*Topics*	*Topics*	*isBefore*
hasLearningContentOf	*LearningObjectives*	*Topics*	*isContentOf*
*isabletoUnderstandKnowledgeOf**			
isabletoApplyKnowledgeOf *			
isabletoAnalyzeProblemOf *			
hasLearningOf	*Courses*	*LearningObjectives*	*isLearnedIn*

*Customized sub-properties for this study's implementation

knowledge of elements among triples will provide graph-based connections which finally define the structure of the curriculum.

Curriculum data in form of a curriculum structure model can easily be transformed into an ontology. Our study adapted the approach of Bussemaker et al. (2017) to map the model into a curriculum ontology. In order to develop a comprehensive method for systematically converting raw curriculum data into curriculum ontology, it is necessary to adjust the methodology of Bussemaker et al. (2017) which was originally developed for manual execution. In composing learning objectives, our modification reduces the sophistication of the learning objective class by substituting the "learning context" entity with a learning verb and learning topic that are feasible for recognition and analysis. This still retains the ability to capture the essence of each learning objective while making it easier for system to accurately detect relevant words in the input text. At the level of knowledge layer, Bussemaker et al. (2017) used the *isA* relation which is equivalent to a subclass, and the *Uses* relation which is equivalent to an object property expressing the associations among the *Topics* instances. These relations are presented in the ontology as different types of ontology components. Our adaptation uses only an object property *continueFrom* to indicate previous topics in the learning sequence, which is applicable to trace the sequence of topics or courses back to the root.

In our adaptation, the high level structure of curriculum ontology is shown in Fig. 4 and the logical associations are summarized in Table 1. The high level classes of the ontology are *Topics*, *LearningObjectives* and *Courses*. The entities in each high level class are defined as subclasses which represent the entities of curriculum objects, namely topic, learning objective and course respectively. The *continueFrom* object property is assigned to create a relation from a topic entity to other(s) in order to convey information of knowledge dependency and continuity. Its backward association is represented by the *isBefore* inverse property.

The *hasLearningContentOf* object property (which the inverse property is *isContentOf*) is a super-property that covers all cognitive levels in logical association from the *LearningObjectives* to *Topics* classes. In general, a sub-property can be customized and assigned under the super-property and will inherit all the super property's characteristics.

The information regarding the learning objectives presented in each course is reflected in the associations between the *Courses* and *LearningObjectives* classes. The *hasLearningOf* object property is utilized for coupling the *Courses* to *LearningObjectives* classes, whereas the *isLearnedIn* inverse property represents the opposite association.

The inverse property enables the ability to track and navigate the connections between objects of the *Topics* and *Courses* classes. These connections are established using the *hasLearningContentOf* and *hasLearningOf* object properties, which are associated with objects of the *LearningObjectives* class. In a similar manner, through *LearningObjectives* class, learning sequences based on the association among the objects of the *Topics* class in the knowledge layer can be referred to by the objects in the courses layer or the learning objectives layer. The relationships within and between layers provide the ability to determine the affiliation and interconnectedness of the object of interest.

The example curriculum data illustrated in Fig. 3 and discussed in "Curriculum structure model" section can be represented using the curriculum ontology as shown in Fig. 5.

Implementation

Sample data

A set of documents called a *qualification framework* was used as source data in our implementation. A qualification framework is a formal system providing an authoritative reference for regional or national education standards (James & Borhene, 2015). The Thai Qualifications Framework for Higher Education (TQF:HEd) was chosen due to the research location and the availability of official documents. The TQF:HEd documents consist of the TQF1 through TQF7 documents. The information which directly involves the subject matter is in the TQF1, TQF2 and TQF3 documents. The National Education Act stipulates that the TQF1 contains the framework of standard educational outcomes for each field of study which all higher education curricula in Thailand must be conformed. The TQF2 is an official document presenting summary information for an individual curriculum maintained by the academic institution. The TQF3 is an official

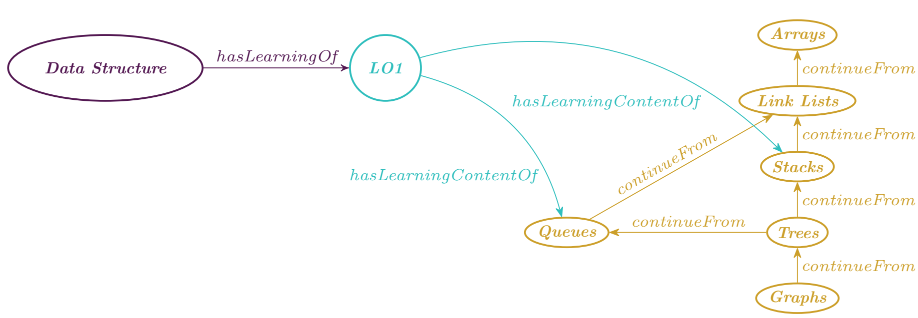

Fig. 5 Example of curriculum ontology

document presenting the learning detail of each individual course that appears in the curriculum. The lecturer responsible for each course is the author of the TQF3 (Ministry of Higher Education, Science, Research and Innovation, 2011; The Government of Thailand, 2020; Ministry of Higher Education, Science, Research and Innovation, 2022).

The implementation in this study used curriculum data from the Bachelor of Science Program in Data Science and Software Innovation (DSSI) and the Bachelor of Science Program in Information and Communication Technology (ICT), Faculty of Science, Ubon Ratchathani University (2020 version) as sample data. The course content was accessed in the form of Portable Document Format (PDF) files, as depicted in Fig. 6 (see full-size figure at https://drive.google.com/file/d/1op1COmHWt3oD8f9Pz9miemhX0FCDFsOI). There was one TQF2 document file per program, plus twenty-four files TQF3 files for the DSSI program and sixteen TQF3 files for the ICT program.

Cognitive levels

The cognitive levels were tailored to align with the certification framework referring to the domains of learning outlined in the TQF1 document. The document specifies three levels of cognitive skills: to understand, to apply and to analyze (Ministry of Higher Education, Science, Research and Innovation, 2011). The cognitive levels customized for this curriculum ontology were respectively assigned as *isabletoUnderstandKnowledgeOf*, *isabletoApplyKnowledgeOf* and *isabletoAnalyzeProblemOf* object properties. As discussed earlier (see Table 1), this established the logical association from the *LearningObjectives* to *Topics* classes inheriting the *hasLearningContentOf* super-property.

Implementation framework

An overall implementation framework was designed to automatically derive curriculum ontology from the sample data. The inputs are TQF2 and TQF3 documents and the output is an ontology in form of Web Ontology Language (OWL) file format. Figure 7 displays the course extraction process, the raw text in the TQF2 file will be parsed and segmented in order to recognize and determine all courses and

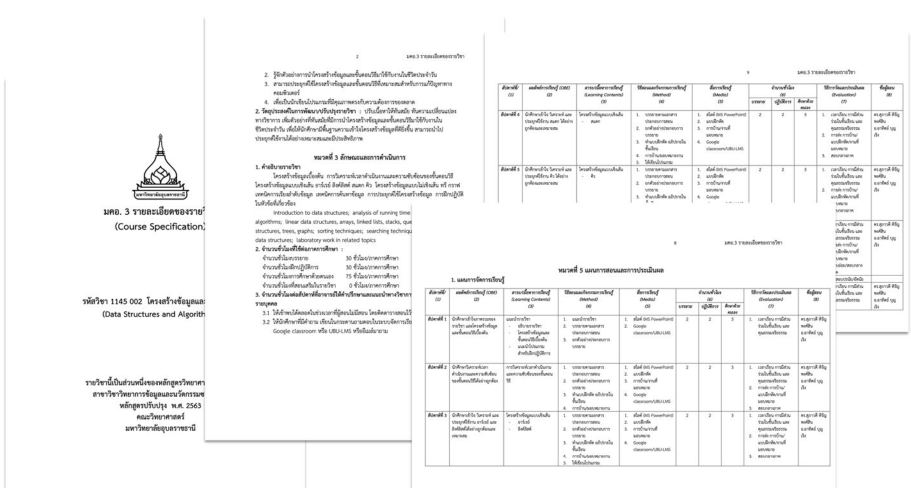

Fig. 6 Example of TQF:HEd documents

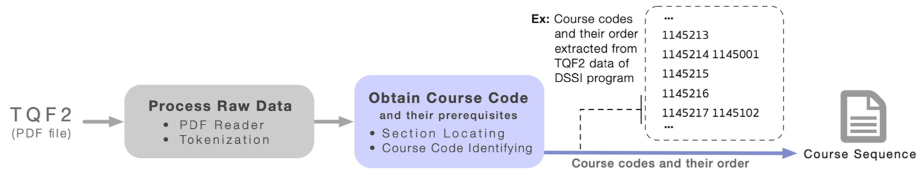

Fig. 7 Course extraction process

their prerequisites, resulting in a compilation of a course sequence list. The course sequence is used in the main ontology construction process.

The main ontology construction process is illustrated in Fig. 8 (see full-size figure at https://drive.google.com/file/d/1vq7S9nA__6SkLo7HXt7aYmnNreynapRr). The raw text in section three and section five of the TQF3 file is parsed and divided into tokens using tokenization function of *pythaiNLP* (Phatthiyaphaibun et al., 2016). Following the regulation of TQF:HEd, section three of the TQF3 outlines all topics covered in each course and section five of TQF3 outlines the syllabus which details the topics and the interrelation between topic and cognitive level. After cleaning and filtering, the tokens from section three and section five are compared one by one. Matched tokens are stored in the ontology as topics. The position of each topic in the syllabus that is the week in which it is supposed to be taught is used to link the sequence of topics. At the same time, the earliest and latest topics are located and stored to be used to construct topic relations between courses. The filtered text tokens of section five from previous process are subjected to the part-of-speech tag to find verb text tokens that may possibly be cognitive levels using POS tagging function of *pythaiNLP* (Phatthiyaphaibun et al., 2016). These verbs are compared with the cognitive level corpus which was predefined from TQF1 information. Matching tokens will be paired with topics from the previous process one by one depending on the found position and stored in the ontology as learning objectives. If there is no tokens that match to the cognitive levels, there will be no learning objective stored in the ontology for that position.

In Fig. 9, the final step is to represent the course sequence at the level of topics. After all TQF3 documents were processed, the relation between courses will be linked by information of the prerequisites and the earliest and latest topics in the course sequence. The process proceeds course by course, identifying the first set of topics, and then conducts a search for the prerequisite course to identify the last set of topics, then links them together. This establishes the connection between courses inside the ontology, based on their component topics. For instance in Fig. 9, the 1145217 course has relations through the *LearningObjectives* class connected to the *IntroductionToSoftwareMeasurement* topic which lead to the *SoftwareMaintanance* topic that belong to the 1145102 course. Using ontology-related functions, these relationships can be investigated in both up and down directions. This ensures the comparable quality, also responds to the design of Curriculum Structure Model.

The ontology created from the DSSI program consists of 24 courses, 170 learning objectives and 204 topics. The ontology created from the ICT program consists of 16 courses, 126 learning objectives and 157 topics. Both ontologies can be viewed online at https://webprotege.stanford.edu/#projects/1b080e48-38e8-44c1-8ffe-cde7b3f5f482 and

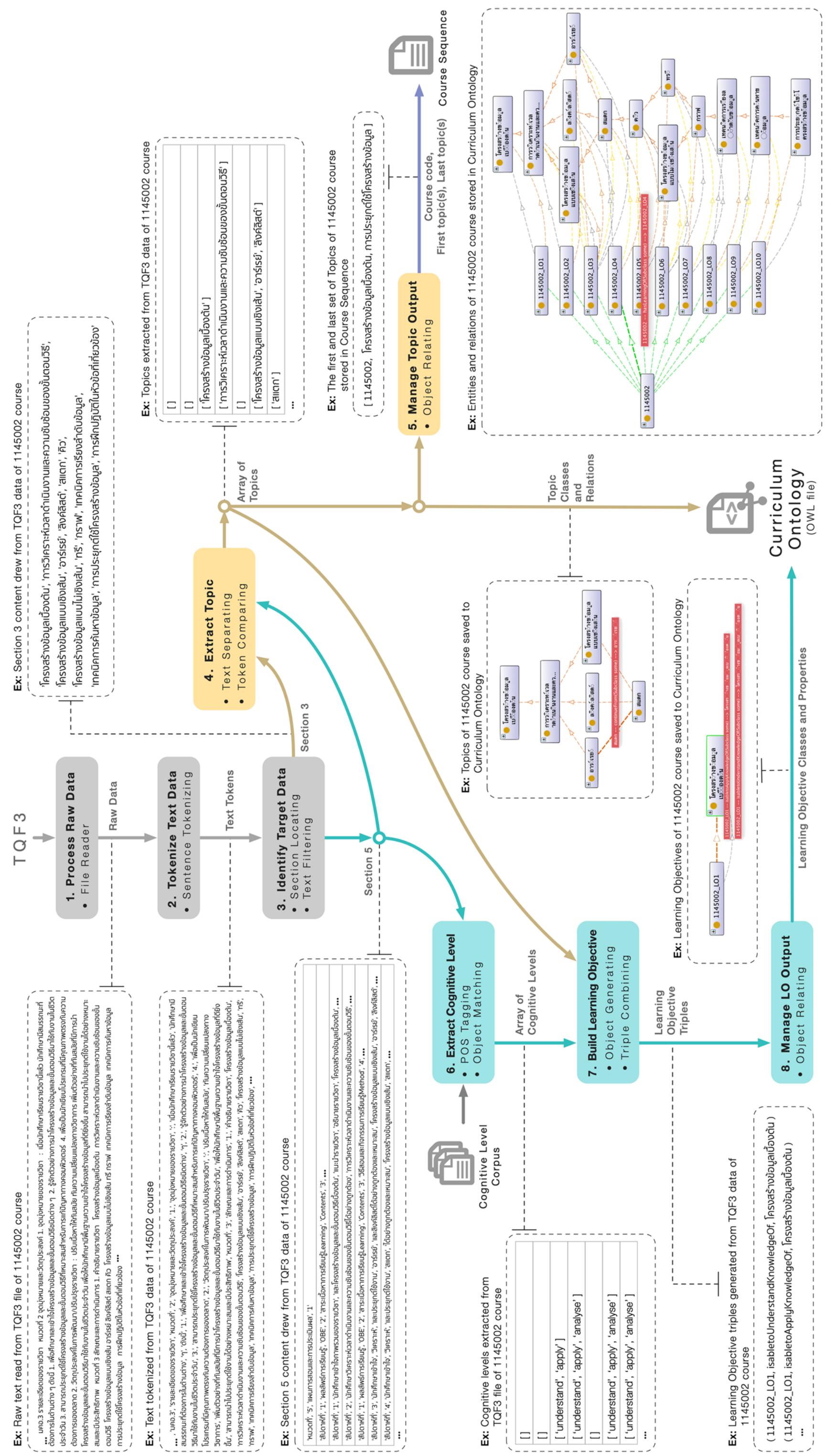

Fig. 8 Main ontology construction process

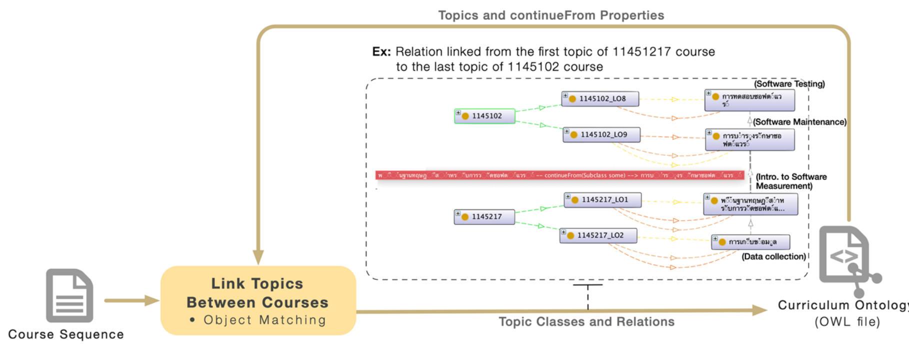

Fig. 9 Course linking process

https://webprotege.stanford.edu/#projects/5de4d366-7166-4a82-851c-63ede1e199d1, respectively.

Evaluation

The evaluation consists of two parts. Part 1 attempts to measure retrieval effectiveness of the implementation, that is, how completely and correctly the process extracted learning topics. Part 2 relies on expert judgement regarding the quality and completeness of the ontology.

Part 1 assessment used precision and recall of the information retrieval performance metrics (Van Rijsbergen, 1981; Ting, 2010). Referring to the TQF3 format, all topics in section three of TQF3 (per one course) are considered as query topics regarding these topics are the topics that supposed to be exist in this course. The possible topics from section five of TQF3 (per one course) are considered as potential retrieved topics as the topics that existed in the ontology were drawn from them. Thus, the precision was calculated by the number of the query topics which appear in the set of retrieved topics (that is, the topics of each course appeared in the curriculum ontology and derived by the implementation) relative to the total number of retrieved topics. The recall was calculated as the number of the query topics which appear in the set of retrieved topics relative to the total number of query topics. Formally, $n(A)$ is the number of all topics that are presented in section three of TQF3, $n(B)$ is the number of possible topics as presented in section five of TQF3, and $n(A \cap B)$ is the number of topics of each course appeared in the curriculum ontology, percentage precision and recall were calculated by these equations:

$$\%Precision = \frac{n(A \cap B)}{n(B)} \times 100 \qquad \%Recall = \frac{n(A \cap B)}{n(A)} \times 100$$

For part 2 assessment, the experts were asked to provide their responses using a structured evaluation form. A separate evaluation form was prepared for each course. Six experts who taught in both programs and were the authors of the TQF3 documents served as evaluators in this assessment. Each assessment form was completed by three experts. There were thirty-two individual courses/forms from the combined DSSI and ICT programs (as there were some common courses). Each form contained different ontology information about topic associations and learning objectives. There were three sections of questions in the form. In section one, the topic associations from the curriculum ontology was prepared as a diagram in picture. The assessor was asked to redraw or correct the diagram based on his/her opinions, using TQF3 document of that course as reference. In section two, there were three questions asking about the accuracy, coverage and sufficiency the ontology information provided, compared to the information in TQF3 document. A five-point scale was applied to this section with five as the highest evaluation (most accurate, etc.) and one as the lowest. Section three provided space for the assessor to enter free-form comments and give feedback.

The results from section one were summarized by comparing the rank of each topic given by the curriculum ontology with the ranks provided by the expert using Spearman's coefficient (r_s) (Spearman, 1904). Rank is order in topic sequence as determined by "continueFrom" relation.

$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$ where d = difference between the rank of topic sequence from the curriculum ontology and the expert revision, n = sample size of topics in sequence. The scores given by experts in section two of the form were summarized using the mean and standard deviation of the Likert ratings for each question.

Results and discussion

In this section, the findings of our study are presented and discussed including the retrieval effectiveness and evaluation results from experts.

Retrieval effectiveness

From part 1 assessment (see "Evaluation"), the results measuring retrieval effectiveness performance of our implementation are presented in Table 2. Collectively, the percentage of precision and recall are 83.14% and 84.84% respectively. These results are favorably compare to the performance measured in the studies of Fiallos & Ochoa (2019) whose research is about semi-automatically generated an ontology for the knowledge layer from existing learning materials. The high value of the precision and recall suggesting our method for deriving *Topic* instances is practical and effective.

Detailed results show considerable variability in precision and recall scores for different courses. Feedbacks from the experts suggested that the poorer scores of some courses are likely due to incomplete document content. In Thailand, the TQF documents are official and compulsory. Every academic institution is required to submit a full set of TQF documents to the regulatory organization before launching any higher education program. In practice, there are many factors that lead to incompleteness and poor quality in these documents, including time pressure, uncertain teaching plans and inconsistent institutional policies. Since we used these documents as the sample data, our results reflect those quality problem. Nevertheless, using TQF documents as input helps narrow down learning content to the exact scope of the course, as well as supporting on implementation that uses relatively simple tools and has low computational demands.

The relationships among topics in the knowledge layer capture vital information required for defining learning paths. The curriculum ontology has a re-classifiable quality whereby the topics can be reassigned to new learning objectives or learning units (such as courses or modules) depending on the needs of learning. Since the relationships between the topics maintain unchanged, it is possible: to trace back to determine

Table 2 Percentage of precision and recall from part 1 assessment

Course code	Course name	Number of topics			Precision	Recall
		Section 3	**Section 5**	**Section 3∩ Section 5**		
1141001	Statistics and Quantitative Analysis	12	13	7	53.85	58.33
1141002	Discrete Mathematics	11	9	9	100.00	81.82
1145000	Introduction to Computer and Computer Organization	14	11	11	100.00	78.57
1145001	Introduction to Programming	10	15	8	53.33	80.00
1145002	Data Structures and Algorithms	14	13	13	100.00	92.86
1145003	System Analysis and Design	7	13	6	46.15	85.71
1145004	Database System	10	16	10	62.50	100.00
1145006	Object-Oriented Programming	7	11	7	63.64	100.00
1145007	Mobile Application Programming	9	10	9	90.00	100.00
1145102	Software Engineering	10	13	9	69.23	90.00
1145103	Software Project Management	9	7	7	100.00	77.78
1145200	Data Warehousing	10	12	9	75.00	90.00
1145202	Natural Language Processing	18	16	13	81.25	72.22
1145203	Decision Support Systems	11	8	8	100.00	72.73
1145205	Data Mining	11	11	9	81.82	81.82
1145207	Pattern Recognition	7	6	6	100.00	85.71
1145208	Deep Learning	6	6	6	100.00	100.00
1145210	Big Data Analytics	10	8	7	87.50	70.00
1145213	Software Testing and Quality Assurance	11	11	11	100.00	100.00
1145215	Requirements Engineering	9	9	8	88.89	88.89
1145216	Software Innovation and Application Development	6	5	5	100.00	83.33
1145217	Software Metrics	9	8	8	100.00	88.89
1145219	C# Programming	12	13	10	76.92	83.33
1145220	Java Programming	11	9	9	100.00	81.82
1146001	Logic Programming	9	9	8	88.89	88.89
1146003	Data Communication and Inter-networking	13	14	12	85.71	92.31
1146004	Innovation and Information Technology Entrepreneurship	11	9	9	100.00	81.82
1146005	Ethical and Legal Issues of Information Technology	10	14	9	64.29	90.00
1146105	Server Platform and Network Administrative	21	17	17	100.00	80.95
1146108	Network Monitoring and troubleshooting	11	10	9	90.00	81.82
1146110	Computer and Network System Maintenance	9	9	9	100.00	100.00
1146111	Electronic Commerce	15	15	13	86.67	86.67
		Average			85.80	85.82
		Standard Deviation			16.95	9.85

which lessons should be learned before, or to define the prerequisites for new learning units, or to provide students self-directed learning paths. A more general and more valid ontology of knowledge layer essentially indicates greater utility of curriculum ontology. However, previous research findings uncovered concerns over the consistency and correctness of the outcome ontologies (Atapattu et al., 2017; Fiallos & Ochoa, 2019). Note that the separation of our model into layers means that we can replace our method for topic extraction with something even more effective, if this becomes available.

Expert feedbacks

The Spearman's coefficient (r_s) values calculated from answers in section one of part 2 assessment form suggest good correspondence between topic structure of the ontology and the experts' views. In Fig. 10, almost all r_s values are above 0.75. The correlation coefficients illustrated in Fig. 10 provided statistically significant evidence of a robust relationship ($\alpha = 0.05$). This can also confirm the correctness of the topic derivation process provided by the implementation.

The pattern of individual item metrics in Table 2 and Fig. 10 shows similar tendencies in many cases. Specifically, courses 1141001, 1145003, 1145219, 1146005 and 1146111 showed poorer performance in both assessments. This consistency strengthens our confidence in the performance measurement. However there are some courses in Table 2 that do not show the same trend in Fig. 10, namely 1145004, 1145006, 1145102, 1145202 and 1145210. According to feedback from the experts in Table 3, documents for these courses were noted as lacking complete information. This defect possibly interfered the ontology extraction, producing incorrect entities or missing relationships.

Detailed assessment results related to learning objectives columns 2 and 4 in Table 3 were sparse. The summary feedback of experts in Table 3 shows very few identified defects in the learning objectives. This may indicate that the learning objective content was mostly viewed as correct. Alternatively, it might suggest that the experts paid less attention to learning objectives than to knowledge topics and their sequence. Since the *Topics* entities are used in composing *LearningObjectives*, the completeness of the *Topics* structure in the knowledge layer also affects the validity of *LearningObjectives*.

Overall assessment scores from the experts regarding the accuracy, coverage and sufficiency of the ontology, presented in Table 4 shows high mean scores with small standard deviations. This implies that in their view, at least, the curriculum ontology

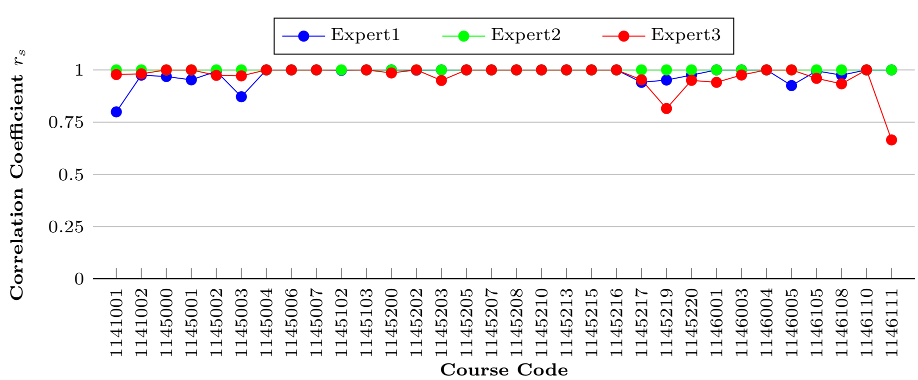

Fig. 10 r_s values between the curriculum ontology and the revision of expert 1–3 given in part 2 assessment

Table 3 Summary of feedback from experts in part 2 assessment

Course code	Course name	Missing some topics	Missing some information of learning objectives	Incorrect order of topics	Incorrect information of learning objectives	Incomplete information in TQF3 document
1141001	Statistics and Quantitative Analysis	✓				✓
1141002	Discrete Mathematics	✓	✓	✓	✓	✓
1145000	Introduction to Computer and Computer Organization	✓		✓		
1145001	Introduction to Programming	✓		✓		✓
1145002	Data Structures and Algorithms	✓		✓		✓
1145003	System Analysis and Design			✓		
1145004	Database System					✓
1145006	Object-Oriented Programming	✓				✓
1145007	Mobile Application Programming					✓
1145102	Software Engineering	✓				✓
1145103	Software Project Management	✓				
1145200	Data Warehousing	✓				
1145202	Natural Language Processing			✓		✓
1145203	Decision Support Systems	✓		✓		
1145205	Data Mining					
1145207	Pattern Recognition					
1145208	Deep Learning					
1145210	Big Data Analytics	✓				✓
1145213	Software Testing and Quality Assurance	✓	✓			
1145215	Requirements Engineering	✓				
1145216	Software Innovation and Application Development					
1145217	Software Metrics	✓		✓		
1145219	C# Programming		✓	✓		✓
1145220	Java Programming	✓		✓		✓

Table 3 (continued)

Course code	Course name	Missing some topics	Missing some information of learning objectives	Incorrect order of topics	Incorrect information of learning objectives	Incomplete information in TQF3 document
1146001	Logic Programming	✓		✓		
1146003	Data Communication and Inter-networking	✓	✓	✓		
1146004	Innovation and Information Technology Entrepreneurship	✓				✓
1146005	Ethical and Legal Issues of Information Technology	✓				✓
1146105	Server Platform and Network Administrative	✓	✓	✓	✓	✓
1146108	Network Monitoring and troubleshooting	✓	✓	✓		✓
1146110	Computer and Network System Maintenance					
1146111	Electronic Commerce	✓		✓		✓

Table 4 Means and standard deviations score of accuracy, coverage and sufficiency from Part 2 assessment

Factor:	Question	Means (SDs)
Accuracy:	*How accurate is the information from the curriculum ontology can present the information in TQF3 documents?*	4.45 (0.44)
Coverage:	*How comprehensive is the information from the curriculum ontology can be obtained from the information in TQF3 documents?*	4.18 (0.51)
Sufficiency:	*How sufficient is the information from the curriculum ontology can represent the information in TQF3 documents?*	4.28 (0.47)

was an effective representation of learning objectives and curriculum content. Given the ability to describe core curricular information, the experts suggested that the curriculum ontology can be utilized in a variety of applications that could potentially improve the regular process of curriculum management. For instance, one of the expert evaluators suggested that the ontology could be applied as a tool to verify whether TQF3 documents have appropriate and adequate content. Another judge noted that the ontology might help curriculum designers verify whether there are some missing learning objectives or topics needed to be added to fulfill the program requirements.

Conclusion and future work

This paper presents a generic schema and demonstrates an approach to analyze general curriculum content by considering essential curriculum elements as objects, then mapping the objects with the ontology components to create a curriculum ontology based on systematic, reproducible processes using the generic schema. A key idea is to employ the resulting generic schema to automatically derive curriculum ontology from general curriculum data.

The evaluation results from both assessment methods yielded positive responses and provided compelling evidence that the resulted curriculum ontology possesses adequate quality to represent essential knowledge, skills, and dependencies from actual curriculum.

Our future work will aim at using our curriculum ontology to contribute to the solution of advanced curriculum manipulation. Ontology representation facilitates reclassification and inference. It's expected to be an alternative approach to module-based curriculum development where our automated processes are extendable to generate the higher-level curriculum units as modules, to reduce manual labor and improve the process of higher education curriculum management.

Abbreviations

DSSI	Data Science and Software Innovation
ICT	Information and Communication Technology
OWL	Web Ontology Language
PDF	Portable Document Format
POS	Part of Speech
TQF	Thai Qualifications Framework
TQF:HEd	Thai Qualifications Framework for Higher Education

Author contributions

P. Piriyapongpipat: Methodology, investigation, funding acquisition, and writing—original draft N. Ditcharoen: Conceptualization, supervision, resources, and funding acquisition S. Goldin: validation, and writing—review & editing.

Funding

This research and innovation activity is funded by National Research Council of Thailand (NCRT) in the fiscal year 2021 Project ID: 125433.

Availability of data and materials

Not applicable.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 4 February 2024 Accepted: 5 May 2024
Published online: 15 May 2024

References

Atapattu, T., Falkner, K., & Falkner, N. (2017). A comprehensive text analysis of lecture slides to generate concept maps. *Computers & Education, 115*, 96–113. https://doi.org/10.1016/j.compedu.2017.08.001

Ávila, L. V., Leal Filho, W., Brandli, L., Macgregor, C. J., Molthan-Hill, P., Özuyar, P. G., & Moreira, R. M. (2017). Barriers to innovation and sustainability at universities around the world. *Journal of Cleaner Production, 164*, 1268–1278. https://doi.org/10.1016/j.jclepro.2017.07.025

Barrier, J., Quéré, O., & Vanneuville, R. (2019). The making of curriculum in higher education: Power, knowledge, and teaching practices. *Revue d'anthropologie des Connaissances, 13–1*(1), 33–60. https://doi.org/10.3917/rac.042.0033

Bussemaker, M., Trokanas, N., & Cecelja, F. (2017). An ontological approach to chemical engineering curriculum development. *Computers & Chemical Engineering, 106*, 927–941. https://doi.org/10.1016/j.compchemeng.2017.02.021

Chakraborty, A., Singh, M., & Roy, M. (2016). Barriers in restructuring university curriculum for a sustainable future. *Annual Research Journal of SCMS, 4*, 67–79.

Chi, Y.-L. (2009). Ontology-based curriculum content sequencing system with semantic rules. *Expert Systems with Applications, 36*(4), 7838–7847. https://doi.org/10.1016/j.eswa.2008.11.048

Chung, H., & Jeongmin, K. (2016). An ontological approach for semantic modeling of curriculum and syllabus in higher education. *International Journal of Information and Education Technology, 6*(5), 365–369. https://doi.org/10.7763/IJIET.2016.V6.715

Ferguson, L. M. (1998). Writing learning objectives. *Journal of Nursing Staff Development, 14*(2), 87–94.

Fernández-Breis, J. T., Castellanos-Nieves, D., Hernández-Franco, J., Soler-Segovia, C., Robles-Redondo, M. D. C., González-Martínez, R., & Prendes-Espinosa, M. P. (2012). A semantic platform for the management of the educative curriculum. *Expert Systems with Applications, 39*(5), 6011–6019. https://doi.org/10.1016/j.eswa.2011.11.123

Hicks, O. (2018). Curriculum in higher education: Confusion, complexity and currency. *HERDSA Review of Higher Education, 5*, 5–30.

Human-Hendricks, N. E., & Meier, C. (2020). A critical review of the role of responsive curricula in optimising learning in higher education. *South African Journal of Higher Education, 34*(4), 77–94. https://doi.org/10.20853/34-4-3522

Kennedy, D., Hyland, A. L., & Ryan, N. (2009). Learning outcomes and competencies. *Using Learning Outcomes: Best of the Bologna Handbook, 33*(1), 59–76.

Komenda, M., Schwarz, D., Švancara, J., Vaitsis, C., Zary, N., & Dušek, L. (2015). Practical use of medical terminology in curriculum mapping. *Computers in Biology and Medicine, 63*(74), 82. https://doi.org/10.1016/j.compbiomed.2015.05.006

Lightfoot, J. M. (2006). Modular curriculum design using personal learning plans and reusable learning components. *Communications of the IIMA, 6*(4), 65–80. https://doi.org/10.58729/1941-6687.1339

Nuntawong, C., Namahoot, C. S., & Brückner, M. (2017). HOME: Hybrid Ontology Mapping evaluation tool for computer science curricula. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 9*(2–3), 61–65.

Oliver, R., Kersten, H., Vinkka-Puhakka, H., Alpasan, G., Bearn, D., Cema, I., & White, D. (2008). Curriculum structure: Principles and strategy. *European Journal of Dental Education, 12*, 74–84. https://doi.org/10.1111/j.1600-0579.2007.00482.x

Peckham, J., & Maryanski, F. (1988). Semantic data models. *ACM Computing Surveys, 20*(3), 153–189. https://doi.org/10.1145/62061.62062

Sand, O., Davis, D., Lammel, R., & Stone, T. (1960). Components of the curriculum. *Review of Educational Research, 30*(3), 226–245. https://doi.org/10.3102/00346543030003226

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15*(1), 72–101. https://doi.org/10.2307/1412159

Vreuls, J., Koeslag-Kreunen, M., van der Klink, M., Nieuwenhuis, L., & Boshuizen, H. (2022). Responsive curriculum development for professional education: Different teams, different tales. *The Curriculum Journal*. https://doi.org/10.1002/curj.155

Andrade, M. S. (2018). A responsive higher education curriculum: Change and disruptive innovation. In *Innovations in higher education* (pp. 1–16). IntechOpen. https://doi.org/10.5772/intechopen.80443

Bakhshi, H., Downing, J. M., Osborne, M. A., & Schneider, P. (2017). *The future of skills: Employment in 2030*. https://futureskills.pearson.com

Burtscher, M., Peng, W., Qasem, A., Shi, H., Tamir, D., & Thiry, H. (2015). A module-based approach to adopting the 2013 ACM curricular recommendations on parallel computing. In *Proceedings of the 46th ACM technical symposium on computer science education* (pp. 36–41). Association for Computing Machinery. https://doi.org/10.1145/2676723.2677270

Cheng, Y., & Nunes, B. P. (2022). The use of semantic technologies in computer science curriculum: A systematic review. arXiv Retrieved from https://doi.org/10.48550/arXiv.2205.00462

Costandius, E., & Bitzer, E. (2015). Curriculum challenges in higher education. In *Engaging Higher Education Curricula: A critical citizenship education perspective* (pp. 9–26). African Sun Media. https://doi.org/10.18820/9781920689698/01

Dehne, J., & Kiy, A. (2019). Using an ontology based competence database for curriculum alignment of portfolio based learning. In *Proceedings of DELFI workshops 2019* (pp. 121–128). Gesellschaft für Informatik e.V.z. https://doi.org/10.18420/delfi2019-ws-114

Doignon, J.-P. (2014). Learning spaces, and how to build them. In *Formal concept analysis* (pp. 1–14). Springer. https://doi.org/10.1007/978-3-319-07248-7_1

Fiallos, A., & Ochoa, X. (2019). Semi-automatic generation of intelligent curricula to facilitate learning analytics. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 46–50). ACM. https://doi.org/10.1145/3303772.3303834

Heggart, K. (2022). Responsive online course design: Microcredentials and non-linear pathways in higher education. In *Global Perspectives on Educational Innovations for Emergency Situations* (pp. 295–303). Springer. https://doi.org/10.1007/978-3-030-99634-5_29

James, K., & Borhene, C. (2015). Dimension 4: Qualifications and qualifications frameworks. In *Level-setting and recognition of learning outcomes: The use of level descriptors in the twenty-first century* (pp. 76–93). UNESCO Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000242887

Kelly, A. (2004). Knowledge and the curriculum. In *The curriculum theory and practice* (5th edn., pp. 25–43). SAGE Publications.

Liu, D., Mikroyannidi, E., & Lee, R. (2014). Semantic web technologies supporting the BBC knowledge & learning beta online pages. In *Proceedings of the linked learning meets LinkedUp workshop: Learning and education with the web of data co-located with 13th international semantic web conference (ISWC 2014)* (Vol. 1254). http://ceur-ws.org/Vol-1254/1 liu.pdf

Lu, W., & Zhang, J. (2011). Construction of curriculum ontology aiming at educational service support: An ontology approach for knowledge-intensive service systems. In *2011 International joint conference on service sciences* (pp. 315–318). https://doi.org/10.1109/IJCSS.2011.70

Mark, M. (2022). The role of microcredentials in modular learning. Techenical Report. The Lifelong Education Commission. https://www.lifelongeducation.uk/research-1/the-role-of-microcredentials-in-modular-learning

Ministry of Higher Education, Science, Research and Innovation. (2011). *The regulations, announcements, and resolutions of the Ministry of Education and the Higher Education Commission from January 2009 to March 2011*. Office of the Higher Education Commission. https://www.ops.go.th/th/ohec-docs/archive-documents/download/269/6362/16

Ministry of Higher Education, Science, Research and Innovation. (2022). *Ministerial regulations of Thai qualification framework for higher education year 2022*. https://www.mhesi.go.th/index.php/en/all-legal/76-ministerial-regulation/7213-2565-5.html

Noy, N.F., & Mcguinness, D. (2001). Ontology development 101: A guide to creating your first ontology. https://protege.stanford.edu/publications/ontologydevelopment/ontology101-noy-mcguinness.html

OECD. (2020). What types of challenges do countries/jurisdictions face in addressing curriculum overload, and what strategies do they use to address these challenges? In *Curriculum overload: A way forward* (pp. 67–84). OECD Publishing. https://doi.org/10.1787/3081ceca-en

O'Malley, B., & Warden, R. (2022). *Global roadmap to 2030 for higher education announced*. https://www.universityworldnews.com/post.php?story=20220525095453173

Phatthiyaphaibun, W., Chaovavanich, K., Polpanumas, C., Suriyawongkul, A., Lowphansirikul, L., & Chormai, P. (2016). *PyThaiNLP: Thai Natural Language Processing in Python, Python Library*. Retrieved from https://github.com/PyThaiNLP/pythainlp

Robinson, S., Arbez, G., Birta, L. G., Tolk, A., & Wagner, G. (2015). Conceptual modeling: Definition, purpose and benefits. In *2015 Winter simulation conference (WSC)* (pp. 2812–2826). IEEE. Retrieved from https://doi.org/10.1109/WSC.2015.7408386

The Government of Thailand. (2020). Policy and strategy of Thailand HESI 2020–2027 and Thailand SRI plan 2020–2022. Retrieved from https://www.mhesi.go.th/index.php/stg-policy/930-2563-2570.html

Ting, K. M. (2010). Precision and Recall. In *Encyclopedia of machine learning* (pp. 781–781). Springer. https://doi.org/10.1007/978-0-387-30164-8_652

Tractenberg, R. E., Lindvall, J. M., Attwood, T., & Via, A. (2020). Guidelines for curriculum and course development in higher education and training. SocArXiv, 1–18. https://doi.org/10.31235/osf.io/7qeht

UNESCO. (2022). Right to higher education: Unpacking the international normative framework in light of current trends and challenges. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000382335

Van Rijsbergen, C. J. (1981). Retrieval effectiveness. In *Information retrieval experiment* (pp. 32–43). Butterworths.

World Economic Forum. (2022). *Catalysing education 4.0: Investing in the future of learning for a human-centric recovery*. Retrieved from https://www.weforum.org/publications/catalysing-education-4-0-investing-in-the-future-of-learning-for-a-human-centric-recovery/

Zouri, M., & Ferworn, A. (2021). An ontology-based approach for curriculum mapping in higher education. In *2021 IEEE 11th Annual Computing and communication workshop and conference (CCWC)* (pp. 0141–0147). https://doi.org/10.1109/CCWC51732.2021.9376163

Publisher's Note