

An Application of Structural Equation Modeling for Developing Good Teaching Characteristics Ontology

Somjin PHIAKOKSONG, Suphakit NIWATTANAKUL,
Thara ANGSKUN

*School of Information Technology, Institute of Social Technology
Suranaree University of Technology, Nakhon Ratchasima, Thailand
e-mail: {somjin, suphakit, angskun}@sut.ac.th*

Received: April 2013

Abstract. Ontology is a knowledge representation technique which aims to make knowledge explicit by defining the core concepts and their relationships. The Structural Equation Modeling (SEM) is a statistical technique which aims to explore the core factors from empirical data and estimates the relationship between these factors. This article presents an application of Structural Equation Modeling to construct a lightweight ontology of good teaching characteristics. The principal factors from SEM are used as concepts of the proposed ontology. The OAsys approach is used as an ontology development process. The proposed ontology consists of three levels: super-class level consists of 6 components of teaching; the first-level sub-class consists of 12 latent factors from SEM; and the second-level sub-class consists of 66 classes of questionnaire items. The experimental result indicated that the proposed ontology which constructs from the core factors of SEM is well-structured, free from a node with multiple parent classes. There is no cycles loop between sub-classes. Having well-balanced of ratio between parent classes and sub-classes. The structure of this proposed ontology corresponds to the principle of conceptual balance.

Keywords: structural equation modeling, good teaching characteristics, ontology development process, lightweight ontology.

1. Introduction

The foundation of a good society is based on the good quality of education of citizens. Education is a process to provide the experience that has a formative effect on the mind, character, or physical ability of an individual, also known as the “teaching and learning process”. Administering the teaching and learning process to achieve high quality of learners consists of several components including teachers, content, learners, environment and assessment (Dees *et al.*, 2007).

The teacher is the most important factor in the teaching process e.g., teachers have a role as content provider and assignment preparation, assessment learning, classroom management, etc. The achievement of learners was directly affected by the quality of

teaching. Hence, to obtain excellence in teaching, teachers should know the characteristics of good teaching, be concerned with the different learning styles of learners and perceive their own strengths and weaknesses in teaching to improve their teaching quality.

To indicate the quality of teaching, an important type of course evaluation called “Student Evaluation Teaching (SET)” surveying is the most widely used in higher education (Moss and Hendry, 2002). A type of students’ feedback that is obtained from SET surveying is answering open-ended questions. These feedbacks are useful for guiding teachers to improve their teaching process and reflect other metric of quality of educational institutes that should not be ignored (Jordan, 2011; Abd-Elrahman *et al.*, 2010). Moreover, these feedbacks can become valuable asset of an educational institute, if it is analyzed systematically.

Recently, Student Evaluation Teaching in electronics based surveying replaced the traditional paper-based (Moss and Hendry, 2002). This phenomenon leads to increased amounts of feedback data which is obstacle to analysis by humans. To overcome this problem, a field of data mining called “Opinion Mining (OM)” was commonly used (Pang and Lee, 2008). There are various fields that implement OM to analyze data which was written as natural language e.g. Legal blog (Conrad and Schilder, 2007), Educational feedback (Abd-Elrahman *et al.*, 2010), Governmental Decisions (Stylios *et al.*, 2010), etc. These have shown that OM can be adopted as an automatic process for analysis of massive feedback data.

However, most of the previous work that dealt with the educational data only focused on data mining technique development. The results that were obtained from these OM processes are expressed in several aspects. For example, the teaching features that were extracted by Abd-Elrahman *et al.* (2010), El-Halees (2011) and Leong *et al.* (2012) are presented in the different aspects of teaching features. These results did not guarantee suitable characteristics that teachers should be aware of. In order to get the proper knowledge, the good teaching characteristics that corresponded with the learning context of educational institute should be explicit declared.

Nowadays, a knowledge representation technique called “Ontology” has become popular. Ontology aims to make knowledge explicit by expressing knowledge structure, capturing knowledge of a certain area by providing relevant concepts and relations between them (Brank *et al.*, 2005). Ontology is used in several areas. It might serve as a backend technology, facilitating the solution of some problems (Rene Robin and Uma, 2011). It can apply to the teaching and learning context in order to structure the subject domain of interest (Boyce and Pahl, 2007).

To develop ontology, identifying concepts by the domain expert who has knowledge in interest domains is required. Presently, there is a statistical approach called “Structural Equation Modeling (SEM)”. This technique aims to explore the core factors from empirical data and also estimates the relationship between these factors. Therefore, this article proposes an application of SEM to explore the core factors that affect the quality of the teaching process. Then, use these core factors to construct ontology of good teaching characteristics.

The structure of this article is organized as follows. Good teaching characteristics, structural equation modeling approach, types of ontology and ontology development processes are introduced in Section 2. In Section 3 explains the methodology to identify and

develop good teaching characteristics ontology. The experimental results of this study are shown in Section 4. The conclusion and future work are presented in Section 5.

2. Literature Review

2.1. Good Teaching Characteristics

Good teaching characteristics depend on various factors such as the content and activity, the experiences that encourage the student’s skill improvement, etc. Knowing about the characteristic and cause of good teaching would be a guideline for a teacher to meet the high quality of teaching in practice. Several educational researchers have studied and proposed the characteristics of good teaching as shown in Table 1.

The previous works have shown there are many items of good teaching characteristics which depend on the different context of studies. However, these characteristics items

Table 1
Summary of previous studies on the good teaching characteristics

Order	Number of good teaching characteristics	Component of good teaching					
		Knowledge	Preparation	Teaching technique	Assessment	Materials	Personality
1	Cooper and Foy (1967)	43			•	•	•
2	Eble (1971)	5	•	•	•	•	•
3	Sheffield (1974)	10	•	•	•		•
4	Ebro (1977)	9	•		•		•
5	Lewis (1982)	8	•		•		•
6	Landbeck (1997)	3			•		•
7	Smith (1980)	8			•	•	
8	Jaitiang (2546: In Thai)	13		•	•		
9	Thompson <i>et al.</i> , (2004)	12		•			•
10	STOW on the wold primary school (2005)	11	•	•		•	•
11	Gurney (2007)	5	•		•	•	•
12	Jahangiri and Mucciolo (2008)	21	•	•	•		•
13	College of Agricultural and Life Sciences, University of Florida (2009)	5	•		•		•
14	Biostatistics, Johns Hopkins University (2009)	39		•	•	•	•
15	Aregbeyen (2010)	17					•
16	Al-hebaishi (2010)	4	•	•	•	•	•

can be grouped into six components: knowledge, preparation, teaching technique, assessment, materials, and personality.

- 1) Knowledge: teacher has enough content knowledge for teaching and answering the questions of students.
- 2) Preparation: teacher has good teaching preparation (contents, process, and materials) before actually teaching.
- 3) Teaching technique: teacher has methods and techniques to transfer their knowledge to the students and also have the ability to control their students in the classroom.
- 4) Assessment: teacher has fair judgment and validity of the assessment process to indicate achievements of students.
- 5) Material: teacher utilizes suitable teaching materials and have assistant teacher to support their teaching process.
- 6) Personality: teacher has good personal behavior and good human relations.

Additionally, most of the previous works indicated that the teaching techniques and personality components are the most important components of good teaching characteristics.

2.2. Structural Equation Modeling

Structural equation modeling (SEM) is a research approach used in many academic disciplines, including information systems and marketing (Jacobson *et al.*, 2009). SEM is a general term that describes a large number of statistical models which are used to test and validate substantive theories with empirical data (Lei and Wu, 2007). This technique combines a measurement model (or Confirmatory Factor Analysis (CFA)) and structural model into a simultaneous statistical test. The patterns of relationships between these latent variables are constructed based on the study of educational theory. SEM is a statistical method to model the relationships among multiple predictor and criterion variables (Hoe, 2008).

Lei and Wu (2007) explained that SEM involves several statistical techniques e.g., Factor analysis, Path analysis and Regression. These statistics are used to evaluate two models: a measurement model and a path model.

- 1) Measurement model: is a measuring of latent variables originated from psychometric theories. Unobserved latent variables cannot be measured directly but are indicated by responses to a number of observable variables (indicators). In social sciences, latent constructs are a set of indirect observation variables (latent variables) such as intelligence or reading ability. These variables and their relationships are often gauged by responses to a battery of items that are designed to tap those constructs. Responses of a study participant to those items are supposed to reflect where the participant stands on the latent variable. Statistical techniques such as factor analysis, exploratory or confirmatory, have been widely used to extract the number of latent constructs underlying the observed responses and to evaluate the

adequacy of each item or variables as indicators for the latent constructs they are supposed to measure.

- 2) Path model: is a statistical approach which is an extension of multiple regressions. It involves various multiple regression models that are estimated simultaneously. This provides a more effective and direct way of modeling mediation, indirect effects, and other complex relationships among variables. Path analysis can be considered a special case of SEM in which structural relations among observed (vs. latent) variables are modeled. Structural relations are hypotheses about directional influences or causal relations of multiple variables (e.g., how the independent variables affect dependent variables). Hence, path analysis (and also the more generalized; SEM) is sometimes referred to as causal modeling. Because analyzing interrelations among variables is a major part of SEM and these interrelations are hypothesized to generate specific observed covariance (or correlation) patterns among the variables, SEM is also sometimes called covariance structure analysis.

SEM has become a popular technique that is used to estimate the relationship between latent variables of an interesting domain. In this work, these latent variables that are obtained from SEM can be considered as a set of core concepts which are used to develop good teaching characteristics ontology.

2.3. *Types of ontologies*

Ontology defines the terms used to describe and represent an area of knowledge. Ontology is used by people, databases and applications that need to share domain information. Ontology ranges from the formal structure (Formal “IS-A” relationship) up to the logical constraints structure (Disjointness, Inverse, Part-of, etc.) (McGuinness, 2003). The simple structures of ontology are related to three kinds of objects including 1) Classes: general things (Concept) in the many domains of interest. 2) Relationships: existing relationships between Classes and 3) Properties: attributes of Classes that they may have (Boyce and Pahl, 2007). There are several types of ontology which are classified based on the characteristic of the ontology components. Ontoware (quote in Rousey, 2005) has proposed three aspects to classify ontology including:

- 1) Classification according to the Purpose:
 - *Application Ontology*: Used in a specific application implementing on an ontology-based reasoning. The typical trade-off between expressiveness and decidability requires limited representation formalism.
 - *Reference Ontology*: Used during development time of applications for mutual understanding and explanation between (human or artificial) agents belonging to different communities, for establishing consensus in a community that needs to adopt a new term or simply for explaining the meaning of a term to someone new to the community.
- 2) Classification according to Specificity:
 - *Generic Ontology*: The concepts defined by this layer are considered to be generic across many fields. Typically, generic ontologies (synonyms are “upper

level” or “top-level” ontology) define concepts such as state, event, process, action, component, etc.

- *Core Ontology*: Core ontologies define concepts which are generic across a set of domains. Therefore, they are situated in between the two extremes of generic and domain ontologies. The borderline between generic and core ontologies is not clearly defined because there is no exhaustive enumeration of fields and their conceptualizations. However, the distinction is intuitively meaningful and useful for building libraries.
- *Domain Ontology*: Domain ontologies express conceptualizations that are specified for a specific universe of discourse. The concepts in domain ontologies are often defined as specializations of concepts in the generic and core ontologies. The borderline between core and domain ontologies is not clearly defined because core ontologies tend to be generic within a domain. Thus, it is usually hard to make a clear-cut decision between generic and core as well as between core and domain ontologies. A concept such as “software component”, “biology of life science”, etc. These ontologies would be placed in the application servers that can be reused in all domain ontologies.

3) Classification according to Expressiveness:

- *Heavyweight Ontology*: Heavyweight ontologies are extensively axiomatized and thus represent an ontological commitment explicitly. The purpose of the axiomatization is to exclude terminological and conceptual ambiguities, due to unintended interpretations. Any heavyweight ontology can have a lightweight version. Many domain ontologies are heavyweight because they should support heavy reasoning (e.g., for integrating database schemata, or to drive complex corporate applications).
- *Lightweight Ontology*: Lightweight ontologies are simple taxonomic structures of primitive or composite terms together with associated definitions. They are hardly axiomatized as the intended meaning of the terms used by the community is more or less known in advance by all members and the ontology can be limited to those structural relationships among terms that are considered relevant.

2.4. Ontology development processes

In the last two decades, several approaches of ontology development were proposed e.g. Enterprise Model Approach, Ontology Development 101, Toronto Virtual Enterprise (TOVE), Methontology, etc. (Fernández-López, 1999; Jones *et al.*, 1998). The different steps of each approach are follow:

Uschold and Gruninger (1996) developed an Enterprise model comprises of four main phases: 1) Identifying a purpose and scope; 2) Building the ontology: ontology capture, ontology coding, integrating existing ontologies; 3) Evaluation: verification and validation; and 4) Documentation.

Noy and McGuinness (2001) proposed seven steps of Ontology Development 101: 1) Determine the domain and scope of the ontology; 2) Consider reusing existing ontologies; 3) Enumerate important terms in the ontology; 4) Define the classes and the class

hierarchy; 5) Define the properties of classes; 6) Define constraint and facet of slot; and 7) Create instances.

Gavrilova *et al.*, (2005) proposed a five-step approach including: 1) Glossary development; 2) Laddering; 3) Disintegration; 4) Categorization; and 5) Refinement.

The last but not the least, an ontology development process called “OASys” which has simplified steps and appropriated in practice to build ontology from scratch was proposed by Bermejo (2007). This approach comprises six stages as shown in Fig. 1.

- 1) Enumerate important terms: this stage is trying to make a list of noun and verb. List of terms (noun and verb) was used as concepts, attributes or relationship to construct ontology.
- 2) Define concept taxonomies: this stage is trying to build a hierarchy structure from a list of terms. Any term that cannot construct the next level of sub-hierarchy may be considered as attributes or relationships.
- 3) Define relations: list of verbs was used to be a relationship between concepts.
- 4) Define attributes: some of the nouns in the list could be considered as attributes of a concept. These attributes usually appear without sub-hierarchy.
- 5) Define instances: the real things from the data were inserted as an instance of a class (also called “individual”).
- 6) Define axioms, rules, and functions: this stage defines any constraints of class in the form of axioms or rules. However, this step depends on characteristics of each class which may not declare any axioms and rules as constraints before described instances.

To obtain the principle structure of good teaching characteristics ontology, application of Structural Equation Modeling in the ontology development process is described in the next section.

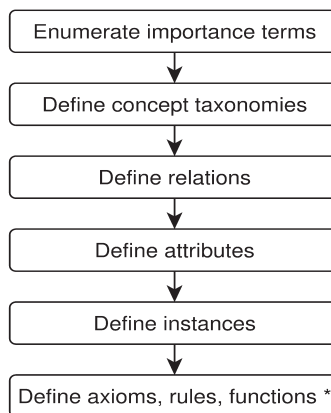


Fig. 1. OASys ontology development tasks.

3. Methodology

This article uses an SEM approach for finding of core factors and then these core factors are used as domain concepts of ontology. The methodology of this work was divided into two main phases including: *Phase 1: Finding good teaching factors*, this phase aims to identify core factors of good teaching characteristics that correspond with the perspective of teacher and learner and *Phase 2: Developing Ontology*, this phase constructs a light-weight ontology of good teaching characteristics. In this phase, the OASys approach is selected as an ontology development process.

3.1. *Finding good teaching factors (Phase 1)*

This phase aims to indicate and verify the good teaching characteristics that correspond with educational theory. The data about good teaching characteristics were surveyed from Thai student and the full-time instructors. The research methodology can be described as follows:

3.1.1. *Population and Samples*

The population is separated into two groups consisting of 1) Teachers: the full time instructors at Suranaree University of Technology (SUT), Thailand, and 2) Students: the learners who are studying at the undergraduate level at SUT. The table for determining sample size (Krejcie and Morgan, 1970) was used to determine of sample size. The total amount of sample units consisting of 97 teachers and 474 students were selected with the simple random sampling technique.

3.1.2. *Development of Research Instruments and Quality testing*

- 1) Review the educational textbooks, academic papers and results of previous work that related to the teaching and learning process, and then synthesizing the characteristics of the teaching process and good teaching characteristics.
- 2) Synthesizing the list of good teaching characteristics items (reviewing of literature as mentioned in Section 2.1). The questionnaire which consists of 66 items based on the 6 teaching components were constructed (source of the questionnaire is shown in Phiakoksong and Angskun, 2011). The numbers of items of each component are shown as follows: Knowledge (4 items), Preparation (4 items), Teaching technique (28 items), Assessment (8 items), Materials (4 items), and Personality (18 items).
- 3) The questionnaire was designed for finding two objectives including: 1) these question items are the good teaching characteristics that correspond to Thailand's learning context; and 2) these good teaching characteristic items are easy to observe by students and/or easy to practice by teachers.
- 4) Quality testing of research instruments, The Index of Item Objective Congruence (IOC) and The Cronbach's α -coefficient (Cronbach, 1951) was computed. These question items are obtained IOC scores between 0.88 and 1.00, which above the minimum threshold (at 0.50). In aspect of reliability of questionnaire, these questionnaires obtained a high reliability rate at 0.983.

3.1.3. Data analysis

The Structural Equation Modeling (SEM) is used to develop a good teaching characteristics model. The phase of data analysis is separated into two stages: 1) Identifying and selecting the good teaching characteristics that correspond with teaching and learning process and 2) Developing the good teaching characteristics model. The process and statistical methods that are used to analyze data can be described as follows:

- 1) Identifying and selecting items: the Index of Item Objective Congruence method is adopted to analyze the closed-end questions. The items that obtained the IOC score higher than the threshold value (at 0.50) would be identified as characteristics of good teaching where the teacher and student are concerned.
- 2) Developing a good teaching characteristics model: This phase consists of two stages, that is 1) the Exploratory Factor Analysis (EFA) is applied to analyze each component of questionnaires. The result of EFA produces new latent variables which are used in SEM model and 2) the second order Confirmatory Factor Analysis (second order CFA) is computed to indicate the fitness of the model with empirical data. The principal factors and factor loading value that affect the teaching and learning process are presented as the final model.

The statistical results of section 3.1 Finding good teaching factors (Phase 1) are described below:

- 1) *Identifying and selecting items*: The 66 items of closed-end questions with 3 choices of answers (“Yes, certainly (+1)”, “Uncertain (0)”, “Absolutely not (-1)”) are answered by the samples. To indicate these question items are good teaching characteristics, the IOC score is adopted. The value at 0.50 is determined as a threshold. Any items of the questionnaire that are equal or higher are selected as good teaching characteristics. The result of this phase is shown in Table 2.

The results in Table 2 show that the teachers and students indicate all of questionnaire items (66 items) describe about the characteristics of good teaching. The teacher has given the IOC scores between 0.83 and 1.00. Teachers indicate that the knowledge (1.00) and preparation (0.99) are important factors of teaching characteristics. While the student

Table 2
Identifying and selecting good teaching characteristics items

Good teaching components	Number of items	IOC score			
		Teacher		Student	
		Min-Max	Average	Min-Max	Average
Knowledge	4	1.00–1.00	1.00	0.95–0.97	0.96
Preparation	4	0.98–1.00	0.99	0.97–0.98	0.97
Teaching technique	28	0.89–1.00	0.97	0.96–0.99	0.97
Assessment	8	0.96–1.00	0.98	0.95–0.97	0.97
Materials	4	0.83–1.00	0.94	0.97–0.98	0.98
Personality	18	0.94–1.00	0.98	0.97–0.98	0.98
Total	66	0.83–1.00	0.98	0.95–0.99	0.97

has given the IOC scores between 0.95 and 0.99. The personality (0.98) and materials (0.98) are important factors of teaching characteristics in aspects of the student. These questionnaire items are used to develop a good teaching characteristics model.

2) *Developing a good teaching characteristics model*: Afterward, these 66 items with 5 point Likert scales are answered. The Exploratory Factor Analysis (EFA) and the second order Confirmatory Factor Analysis (second order CFA) are used to develop a good teaching characteristics model.

At the first stages, the EFA is computed to extract the principal factors (latent variables) and factor loading scores of each component. These principal factors are assigned as variables in the stages of model development. The factor loading (1st factor loading) of these principal factors is obtained. In the second stages, the second order CFA was used to model the good teaching characteristics from these principal factors. The factor loading of the six core components are obtained (2nd factor loading). The result of this phase is shown in Table 3.

Table 3 showed that 12 factors (1st factors) are expressed in this questionnaire. Four of the six components consists of 1 factor including: Knowledge (KN_FUND: 0.280), Preparation (TE_PREP: 0.350), Assessment (ME_TECH: 0.320) and Material (TE_MATE: 0.450) components. While Teaching technique components consist of 6 factors with factor loading scores between 0.048 and 0.090 (UT_FEEDS: 0.048, ST_CENT: 0.059, PT_KNOW: 0.077, KN_TRANS: 0.083, CL_ADMIN: 0.083 and TE_STRUCT: 0.090). The Personality component consists of 2 factors with factor loading scores of 0.150 and 0.160 (INDI_PER: 0.150 and HU_RELAT: 0.160).

Table 3
Core components and principal factors of good teaching characteristics model

Good teaching components	Principal factors of each components	1 st factor loading	2 nd factor loading
1. Knowledge (<i>KNOWLEDG</i>)	1.1 Knowledge fundamental (<i>KN_FUND</i>)	0.280	2.55
2. Preparation (<i>PREPARE</i>)	2.1 Teaching preparation (<i>TE_PREP</i>)	0.350	2.19
3. Teaching technique (<i>TECHNIQU</i>)	3.1 Knowledge transferring technique (<i>KN_TRANS</i>)	0.083	4.57
	3.2 Classroom administration (<i>CL_ADMIN</i>)	0.083	
	3.3 Utilizing the feedback (<i>UT_FEEDS</i>)	0.048	
	3.4 Practical knowledge transferring technique (<i>PT_KNOW</i>)	0.077	
	3.5 Supporting student-centered learning (<i>ST_CENT</i>)	0.056	
	3.6 Teaching is structured (<i>TE_STRUCT</i>)	0.090	
4. Assessment (<i>ASSESSME</i>)	4.1 Measurement and evaluation Techniques (<i>ME_TECH</i>)	0.320	2.01
5. Material (<i>MATERIAL</i>)	5.1 Teaching material and personnel support (<i>TE_MATE</i>)	0.450	1.75
6. Personality (<i>PERSONAL</i>)	6.1 Human relationship (<i>HU_RELAT</i>)	0.160	3.90
	6.2 Individual personality (<i>INDI_PER</i>)	0.150	

$\chi^2 = 27.77$, $df = 31$, $p\text{-value} = 0.63$, $RMSEA = 0.00$, $GFI = 0.99$, $AGFI = 0.98$, $CFI = 1.00$, $SRMR = 0.019$.

The second order CFA revealed that the Teaching techniques component obtained the highest of the factor loading at 4.57. The next important component is the Personality component with the 3.90 of the factor loading. To indicate model fitting, this model obtains the 27.77 of Chi-Square with 31 degrees of freedom, p-value is 0.63, RMSEA is 0.00, GFI is 0.99, AGFI is 0.98, CFI is 1.00 and SRMR is 0.019. Compared with the threshold values, that is p-value > 0.05, RMSEA < 0.07, GFI > 0.95, AGFI > 0.95, CFI > 0.95 and SRMR < 0.08 (Hooper *et al.*, 2008). These statistical indicators indicate that the good teaching characteristics model is consistent with the empirical data.

In the next step, these core components as described above consisting of the principal factors and the detail of questionnaire items are analyzed according to OASys process. The process of ontology development is described in Phase 2.

3.2. Developing Ontology (Phase 2)

The OASys approach is used as an ontology development process of good teaching characteristics. The step of OASys can be described as below:

- 1) *Enumerate important terms*: the questionnaire from phase 1 is used as a source of important terms. Numerous important terms e.g. knowledge, learner, subject, demonstrate, laboratory, supporting, understanding, utilizing, etc. are extracted. The lists of important terms of questionnaire are comprised of 124 of noun and 63 of verb phrases.
- 2) *Define concept taxonomies*: hierarchical structure of good teaching characteristics was constructed. The latent factors that are obtained from SEM results are used as a class of the ontology. The 6 core components are used as super-class, the 12 principal factors (latent variables) are assigned as the first level of sub-class (1st sub-class). While, the questionnaire items are assigned as second level of sub-class (2nd sub-class). In this step, the important keywords of 66 items are used to assign as the meaningful names. This initial structure of good teaching characteristics ontology is shown in Fig. 2.
- 3) *Define relations*: although, the list of verbs is considered as a relationship between concepts, however, all concepts of the initial ontology are expressed as upper-levels which not depend on the list of verbs. Therefore, “IS-A” relation are assigned as relationship name of each level as shown in Fig. 3. While, three relationships i.e. “has-feedback”, “has-feature” and “has-opinion” are determined as the relationship of the instances which used to support the task of student feedback categorization (in opinion mining process).
- 4) *Define attributes*: list of important terms from step 1 are grouped and used as attributes of the second level of sub-class. This ontology is designed to adopt as a knowledge base to store knowledge from an automatic system (opinion mining process). The two main groups of words called Product features (in this case is “Teaching features”) and Opinion (Popescu and Etzioni, 2005) that corresponds with the task of opinion mining were generated. The product features consist of two groups including: 1) *Object group*: the list of real things that arises in teaching process such as persons, places, sources, behavior and teaching activity; and 2) *Action*

group: list of verb terms which acts in the teaching process. While the opinion consists of only one group, this is 3) *Quality group*: these terms indicate the attitude or satisfaction of students in the teaching process, these can be considered as opinion group. These groups appear in couples form $\{Product\ feature \leftarrow (Object, Actions)\}$,

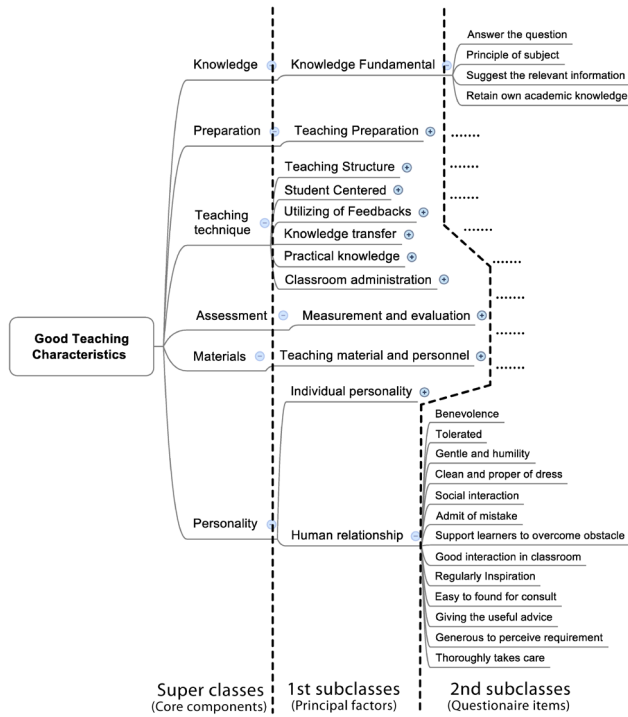


Fig. 2. Initial structure of good teaching characteristics ontology.

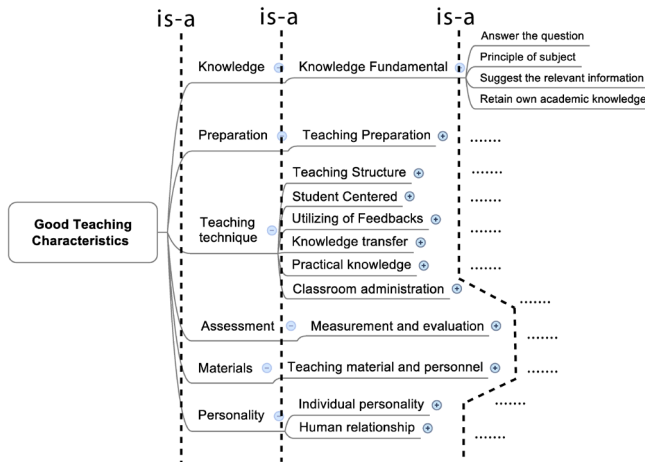


Fig. 3. Relationship of upper-level of good teaching characteristics ontology.

Opinion-(*Quality*)). This couple form is used as the instance (individual) of second level of subclass as shown in Fig. 4.

5) *Define instances*: the descriptions of questionnaire items are used as the source of initial instances. The important terms and their syntactic dependency of question items are extracted as couples form. For example: in Class “Answer the question” has a describe “*Having sufficient knowledge to answer the most of question* from students”. The two initial instances of the example are

- {(<object>*knowledge*</object>, <action>*having*</action>), <quality>*sufficient*</quality>} and
- {(<object>*question*</object>, <action>*answer*</action>), <quality>*most*</quality>}

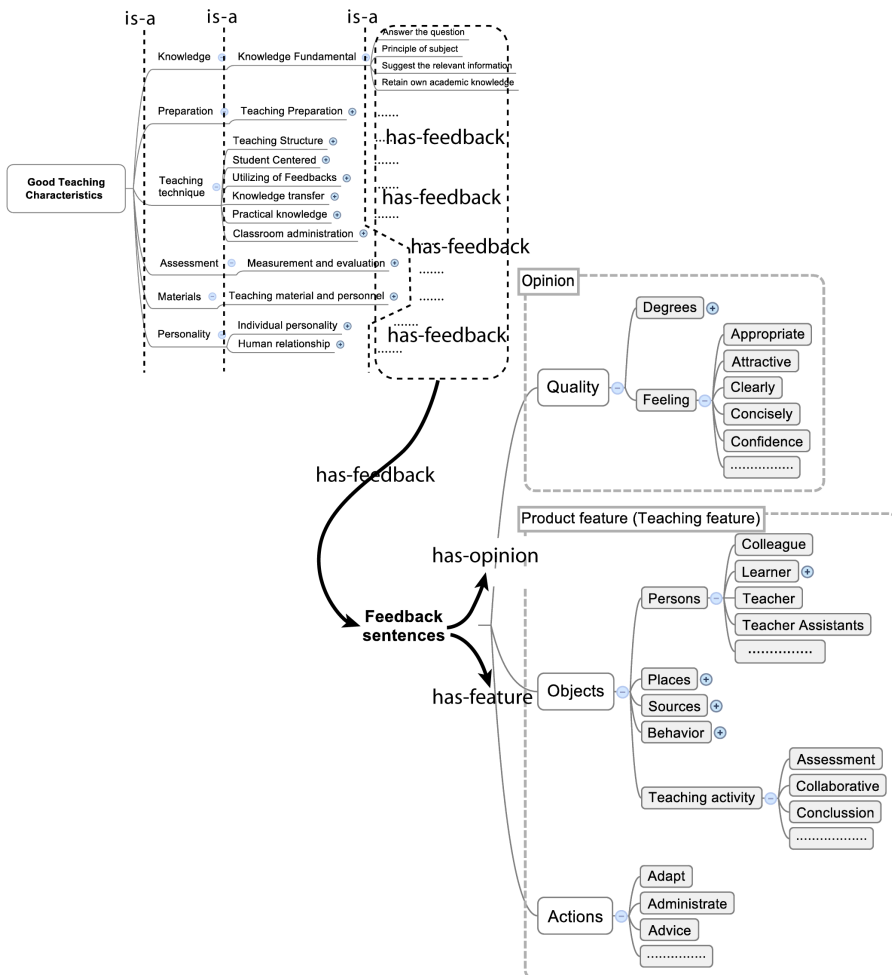


Fig. 4. Attributes of Good teaching characteristics ontology.

6) *Define axioms, rules, and functions*: Although, Zaihrayeu *et al.*, (2007) guided the way to construct the axioms of ontology by utilizing natural language pattern. However, there are several terms appearing in the student feedback. Declaring axioms and strict rules are complicated. Therefore, the explicit declaration of axioms is discarded. However, to overcome this problem, semantic linguistics resource e.g. Wordnet (George, 1995), Asian Wordnet (Sornlertlamvanich, 2010) or specific lexical corpus with the categorization in the opinion mining process is used as an automatic process. This automatic process is used to analysis and classified of these student feedback into the correspond class of the ontology.

According to OASys process as described above, development of this lightweight ontology is implemented with the Protégé Ontology Editor (Protégé, 2003). The experimental result is shown in the next section.

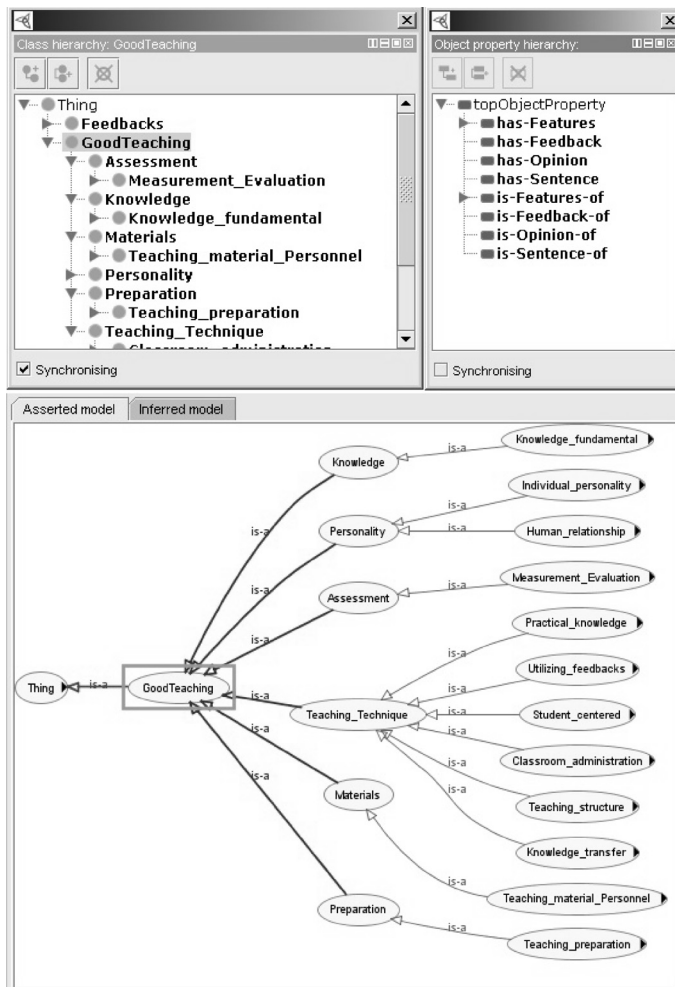


Fig. 5. Lightweight ontology of good teaching characteristics.

4. Experimental Results

4.1 The lightweight ontology of good teaching characteristics

According to the methodology as mentioned above, the structure of the proposed ontology consists of two parts i.e., good teaching characteristics and feedbacks structure. The lightweight ontology of good teaching characteristics is constructed with ontology editor software named “Protégé” (as shown in Fig. 5).

4.2 Ontology Evaluation

Evaluating of the proposed ontology, the Ontology cognitive ergonomics evaluation based on graph topology (Gavrilova *et al.*, 2010a, 2010b; Ontoeval) is implemented to analyze the structure of the ontology. The aim of this evaluation method is to assess ontology balance and its perception by humans. The results are reported in the aspect of ontology structure including the Tangledness and Yngve-Miller metric. Tangledness metrics is the number of vertices of an ontology graph divided by the number of those vertices that have some direct superclasses. This metric is zero for ontologies without multiple inheritance (IS-A relationship). The smaller result value, the ontology is better from the point of view of cognitive ergonomics. Yngve-Miller shows percentage of concepts node and number of links which is hard or easy to perceive by humans. The greater percentage value, the ontology is easy to perceive by humans (Gavrilova *et al.*, 2010a, 2010b).

The evaluation result of the proposed ontology is shown in Table 4.

The results in Table 4 indicated that the proposed ontology is well-structured. There are no subclasses with several parent classes (Nodes with several parent=0). There is no

Table 4
Excerpt of evaluation results of ontology cognitive ergonomics

Evaluation lists	Details
General Metrics	<ul style="list-style-type: none"> • Has cycles = false • Nodes number = 100 • Leaves number = 75 • Max level = 3
Tangledness	<ul style="list-style-type: none"> • Nodes with several parent = 0 • Average Number of Parents Per Node = 0.979 • Tangledness = Infinity // *no nodes with several parents* // • Tangledness Inverse = 0.00
Yngve-Miller metric	<ul style="list-style-type: none"> • Node Degrees = 100 • Node Degrees (except leaves) = 25 • Normal nodes (except leaves) = 23 • Normal nodes (except leaves)/All nodes (without leaves) = 0.92 • 90%-line Node Degree (except leaves) = 7.8 • Max difference from normal degree = 6 • [Yngve-Miller Problems] • <i>Human_relationship = not-normal (degree > 8)</i> • <i>Measurement_Evaluation = not-normal (degree >8)</i>

cycles loop between subclasses (Has cycles=false). Having well-balanced of ratio between parent classes and sub-classes (Normal nodes/ All nodes=0.92) which human are able to perceive. The structure of this proposed ontology corresponds to the principle of conceptual balance (Gavrilova *et al.*, 2010a, 2010b) that is:

- The concept of one level should be linked with the parent concept by one type of relationship such as “is-a”, or “has part”.
- The depth of the branches should be more or less equal (± 2 nodes).
- The general outlay should be symmetrical.
- Cross-links should be avoided as much as possible.

However, the Yngve-Millermetric reported that “*Human_relationship*” and “*Measurement_Evaluation*” are not-normal concept nodes. These two concepts have a number of child nodes more than 8 nodes per classes which meant these classes are not fan-outness characteristics. This problem can be eliminated by restructuring the ontology such as grouping some sub-classes into a concept node or assign sub-classes as parent class. However, Sosnovsky and Gavrilova (2006) indicated that the structure of any ontology can adapt to follow the purpose of use. Therefore, the structure of these classes has retained the same initial structure.

5. Conclusion

This article proposes a technique to develop of good teaching characteristics ontology. The statistical technique called “Structural Equation Modeling” (SEM) was implemented to extract and find the domain concepts and relationship structures of these concepts. Two phases of ontology development were proposed consists of: 1) Finding good teaching factors (Phase 1) and 2) Developing Ontology (Phase 2). In Phase 1, the domain concepts are extracted and verified based on educational theory. This phase uses the Structural Equation Modeling process. The results of this phase are the group of principal factors and their relationships. This phase ensures that these good teaching characteristics correspond with the educational theory and fit with the empirical data. In Phase 2, the OAsys approach is used as the main process to develop ontology of good teaching characteristics. OAsys approach consists of six stages: 1) Enumerate important terms; 2) Define concept taxonomies; 3) Define relations; 4) Define attributes; 5) Define instances; and 6) Define axioms, rules, and functions. Finally, the lightweight ontology of good teaching characteristics which is based on the principal factor from SEM is constructed. The proposed ontology consists of three levels: Super-class consists of 1) Knowledge, 2) Preparation, 3) Teaching technique, 4) Assessment, 5) Material, and 6) Personality; First-level sub-class consists of 1) Knowledge fundamental, 2) Teaching preparation, 3) Knowledge transferring technique, 4) Classroom administration, 5) Utilizing the feedback, 6) Practical knowledge transferring technique, 7) Supporting student-centered learning, 8) Teaching with structured, 9) Measurement and evaluation Techniques, 10) Teaching material and personnel support, 11) Human relationship, and 12) Individual personality; and Second-level sub-class consists of 66 classes of questionnaire items. The ontology cognitive ergonomics evaluation results indicated that this proposed ontology is well-structured.

This ontology is free from a node that has multiple parent classes. There is no cycles loop between sub-classes and being well-balanced between number of parent classes and sub-classes. The structure of the proposed ontology corresponds to the principle of conceptual balance. These evaluation results show that using principal factors from SEM as conceptual nodes is an effective technique to construct ontology. These concept nodes and their relationships are based on the study of educational theory and correspond to the empirical data. This technique is helpful for the ontology developer to express the exact upper structure of ontology. For further development, the proposed ontology could be implemented as a knowledge base to store the knowledge from an automatic system (opinion mining process). Other ontologies that relate to opinion mining such as Marl ontology (Westerski *et al.*, 2011) and KDO ontology (Thalhammer *et al.*, 2011) should be implemented in conjunction with this proposed ontology. In addition, to obtain the heavyweight ontology, linguistic analysis of specific domain (feedback in the educational context) could be explored to construct the axioms of this ontology.

References

- Abd-Elrahman, A., Andreu, M., Abbott, T. (2010). Using text data mining techniques for understanding free-style question answers in course evaluation forms. *Research in Higher Education Journal*, 9, 12–23.
- Al-hebaishi, S. (2010). *Characteristics Of A Good Teacher*.
http://www.dr-safaa-efl.com/wp-content/uploads/2010/01/Char_Good_Teacher1.doc.
- Aregbeyen, O. (2010). Students perceptions of effective teaching and effective lecturer characteristics at the university of ibadan, nigeria. *Pakistan Journal of Social Sciences*, 7(2), 62–69.
- Bermejo, J. (2007). A simplified guide to create an ontology. *The Autonomous Systems Laboratory*, Universidad Politecnica de Madrid.
<http://tierra.aslab.upm.es/documents/controlled/ASLAB-R-2007-004.pdf>
- Brank, J., Grobelnik, M., Mladenić, D. (2005). A survey of ontology evaluation techniques. In: *Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005)*. Ljubljana: Slovenia, 166–170.
- Biostatistics, Johns Hopkins University. (2009). *TA Training Day*.
<http://www.biostat.jhsph.edu/research/TATraining/listing.pdf>
- Boyce, S., Pahl, C. (2007). Developing domain ontologies for course content. *Educational Technology & Society*, 10(3), 275–288.
- College of Agricultural and Life Sciences. (2009). University of Florida, *Characteristics of Good Teaching*.
http://cals.ufl.edu/faculty_staff/pdfs/teachingtips/2009-10-Characteristic-sofGood Teaching.pdf.
- Conrad, J.G., Schilder, F. (2007). Opinion mining in legal blogs. In: *Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL07)*. Palo Alto, CA: USA, 231–236.
- Cooper, B., Foy, J.M. (1967). Evaluating the effectiveness of lectures. *Higher Education Quarterly*, 21, 182–185.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Dees, M.D., Ingram, A., Kovalik, C., Huffman, M.A., McClelland, A., Justice, L. (2007). A transactional model of college teaching. *International Journal of Teaching and Learning in Higher Education*, 19(2), 130–139.
- Eble, K. (1971). *The Recognition and Evaluation of Teaching*. American Association of University Professors, Salt Lake City.
- Ebro, L. (1977). *Instractional Behavior Patterns of Distinguished University Teachers*. Ph.D. Dissertation, Ohio State University.
- El-Halees, A. (2011). Mining opinions in user-generated contents to improve course evaluation. *Software Engineering and Computer Systems: CCIS Part II*, 180, 170–115.
- Fernández-López, M. (1999). Overview of methodologies for building ontologies. In: *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5)*. Stockholm: Sweden.
- Gavrilova, T., Farzan, R., Brusilovsky, P. (2005). One practical algorithm of creating teaching ontologies. In: *Proceeding of Network Based Education*. <http://www.cs.pitt.edu/~rosta/docs/PDF/NBE.pdf>

- Gavrilova, T., Gorovoy, V., Bolotnikova, B. (2010a). Cognitive ergonomics of teaching ontologies. In: *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics (ECCE '10)*. ACM, New York:USA, 271–278.
- Gavrilova, T., Gorovoy, V., Bolotnikova, B. (2010b). Evaluation of the cognitive ergonomics of ontologies on the basis of graph analysis. *Scientific and Technical Information Processing*, 37(6), 398–406.
- George, A.M. (1995). WordNet: a lexical database for english. *Communications of the ACM*, 38(11), 39–41.
- Gurney, P. (2007). Five factors for effective teaching. *New Zealand Journal of Teachers' Work*, 4(2), 89–98.
- Hoe, S.L. (2008). Issues and procedures in adopting structural equation modeling technique. *Journal of Applied Quantitative Method*, 3(1), 79–83.
- Hooper, D., Coughlan, J., Mullen, M.R. (2008). Structural equation modeling: guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, 6(1), 53–60.
- Jacobson, D., Pehlivan, E., Vilovsky, S., Wong, W. (2009). Combining web mining techniques and structural equations modeling for measuring e-commerce perceptions. *Case Studies in Business, Industry and Government Statistics (CS-BIGS)*, 2(2), 99–108.
- Jahangiri, L., Mucciolo, T.W. (2008). Characteristics of effective classroom teachers as identified by students and professionals: a qualitative study. *Journal of Dental Education*, 72(4), 484–493.
- Jaitiang, A. (2546). *Principle of Teaching (3rd ed. revised)*. Bangkok: Odient Store. (In Thai)
- Jordan, D.W. (2011). *Re-thinking Student Written Comments in Course Evaluations: Text Mining Unstructured Data for Program and Institutional Assessment*. Ph.D. Dissertation, California State University, Stanislaus.
- Jones, D., Bench-capon, T., Visser, P. (1998). *Methodologies for Ontology Development*.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.52.2437&rep=rep1&type=pdf>.
- Krejcie, R.V., Morgan, D.W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30, 607–610.
- Landbeck, R. (1997). What students think about good teaching: an exploratory survey at the University of the South Pacific. In: *Proceeding of the Higher Educational Research and Development Society of Australasia (HERDSA) Annual Conference held in Adelaide*. Australia, 17–30.
- Lei, P., Wu, Q. (2007). Introduction to structural equation modeling: issues and practical considerations. *Educational Measurement: Issues and Practice*, 26(3), 33–43.
- Leong, C.K., Lee, Y.H., Mak, W.K. (2012). Mining sentiments in SMS texts for teaching evaluation. *Expert System with Application*, 39, 2548–2589.
- Lewis, K.G. (1982). *The Large Class Analysis Project (Final Report)*. Austin, TX: Texas University Center for Teaching Effectiveness.
- McGuinness, D.L. (2003). Ontologies come of age. In: Fensel, D., Hendler, J., Lieberman, H., Wahlster, W. (Eds.), *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, 171–194.
<http://www-ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-%28with-citation%29.htm>
- Moss, J., Hendry, G. (2002). Use of electronics survey in course evaluation. *British Journal of Educational Technology*, 33(5), 583–592.
- Noy, N.F., McGuinness, D.L. (2001). Ontology development 101: a guide to creating your first ontology. *Stanford Knowledge Systems Laboratory, Technical Report KSL-01-05*.
<http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf>
- Ontoeval, COAT – Cognitive Ontology Assessment. <http://code.google.com/p/ontoeval/>
- Pang, B., Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135.
- Phiakoksong, S., Angskun, T. (2011). Good teaching characteristics of suranaree university of technology: perspective of teacher and learner. *Suranaree Journal of Social Science*, 5(2), 109–130. (In Thai)
- Popescu, A., Etzioni, O. (2005). Extracting product features and opinions from reviews. In: *Proceeding HLT '05 Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*. Stroudsburg, PA: USA, 339–346.
- Protégé. (2003). *The Protégé Ontology Editor and Knowledge Acquisition System*.
<http://protege.stanford.edu/>
- Rene Robin, C.R., Uma, G.V. (2011). Development of educational ontology for software risk analysis. In: *Proceedings of the 2011 International Conference on Communication, Computing & Security (ICCCS '11)*. ACM, New York, USA, 610–615.
- Roussey, C. (2005). *Guidelines to Build Ontologies: A Bibliographic Study, Technical report n°1*. LIRIS, INSA de Lyon.

- <http://www.towntology.net/Documents/guidelines.pdf>.
- Stylios, G., Christodoulakis, D., Besharat, J., Vonitsanou, M., Kotrotsos, I., Koumpouri, A., Stamou, S. (2010). Public opinion mining for governmental decisions. *Electronic Journal of e-Government*, 8(2), 203–214.
- Sheffield, E. (1974). *Teaching in the Universities – No One Way*. Montreal: Mc Gill-Queen’s University Press.
- Smith, R. (1980). *A Checklist for Good Teaching*. Montreal, Quebec, Canada: Teaching and Learning Centre, Concordia University.
- Sornlertlamvanich, V. (2010). Asian WordNet: development and service in collaborative approach. In: *The 5th International Conference of the Global WordNet Association (GWC-2010)*. Mumbai: India.
- Sosnovsky, S., Gavrilova, T. (2006). Development of educational ontology for C-programming. *International Journal “Information Theories & Applications”*, 13, 303–308.
- Stow on the Wold Primary School. (2005). *Characteristics of Effective teaching and learning checklist*.
<http://www.stow-on-the-wold.gloucs.sch.uk/Policies/thepolicies/characteristicsofeffectiveteachinglearning.pdf>
- Thalhammer, A., Toma, I., Hasan, R., Simperl, E., Vrandečić, D. (2011). How to represent knowledge diversity. In: *10th International Semantic Web Conference (ISWC11)*.
http://iswc2011.semanticweb.org/fileadmin/iswc/Papers/PostersDemos/iswc11pd_submission_74.pdf. (Poster)
- Thompson, S., Greer, J.G., Greer, B.B. (2004). Highly qualified for successful teaching: Characteristics every teacher should possess. *Essays in Education*, 10, 1–9.
<http://www.usca.edu/essays/vol102004/thompson.pdf>
- Uchold, M., Gruninger, M. (1996). Ontologies: principles, methods and applications. *Knowledge Engineering Review*, 11(2), 93–155.
- Westerski, A., Iglesias, C.A., Rico, F.T. (2011). Linked opinions: describing sentiments on the structured web of data. In: *Proceeding of 4th international workshop Social Data on the Web (SDoW2011)*. Bonn: Germany.
http://www.gi2mo.org/files/papers/iswc2011/marl_iswc_paper.pdf.
- Zaihrayeu, I., Sun, L., Giunchiglia, F., Pan, W., Ju, Q., Chi, M., Huang, X. (2007). From web directories to ontologies: natural language processing challenges. In: *Proceedings of the 6th International The Semantic Web and 2nd Asian Conference on Asian Semantic Web Conference (ISWC’07/ASWC’07)*, 623–636.

S. Phiakoksong received Bachelor of Education in Measurement and Evaluation from Nakhon Ratchasima Rajabhat University, Thailand, in 1998. He received his Master of Information Science from Suranaree University of Technology, Thailand, in 2010. Currently, he is a Ph.D. Student at School of Information Technology, Suranaree University of Technology, Thailand. His research interests are knowledge management, information retrieval and data mining.

S. Niwattanakul received his Ph.D. in Computer Science from the University of La Rochelle, France in 2008. He is a lecturer at the school of Information Technology, Suranaree University of Technology, Nakhon Ratchasima, Thailand. His current research is about Semantic Web technologies applied to information extraction and retrieval system.

T. Angskun received Bachelor and Master of Engineering in computer engineering from Kasetsart University, Thailand in 1999 and 2002, respectively. He received his Ph.D. in Computer Science from the University of Tennessee, Knoxville, USA in 2007. Currently, he is a lecturer at School of Information Technology, Suranaree University of Technology, Nakhon Ratchasima, Thailand. His major research interests are in parallel and distributed environment, message passing, high performance computing, computer networking, cluster and grid computing.

Struktūrinių lygčių modeliavimo metodo taikymas kuriant gero mokymo charakteristikų ontologiją

Somjin PHIAKOKSONG, Suphakit NIWATTANAKUL, Thara ANGSKUN

Ontologija yra žinių vaizdavimo priemonė, kuria siekiama išreikšti žinias apibrėžiant jų pagrindines sąvokas ir sąryšius. Struktūrinių lygčių modeliavimas (SLM) yra statistinis metodas, leidžiantis iš empirinių duomenų išskirti pagrindinius faktorius ir įvertinti jų sąryšius. Straipsnyje pristatomas šio metodo taikymas kuriant „lengvasvorę“ gero mokymo charakteristikų ontologiją. SLM gauti pagrindiniai faktoriai panaudoti kaip siūlomos ontologijos sąvokos. Ontologijos kūrimo procese pritaikytas „OAsys“ metodas. Siūloma ontologija sudaryta iš trijų lygių: viršutinė klasė sudaryta iš 6 mokymo komponentų, pirmo lygio poklasė sudaryta iš 12 faktorių, gautų iš SLM, antro lygio poklasė sudaryta iš 66 klasių, susidedančių iš klausimyno klausimų. Atlikto eksperimento rezultatai rodo, kad pasiūlyta ontologija sukonstruota iš SLM gautų faktorių yra gerai struktūruota, nėra ciklų tarp poklasių, gerai subalansuotas santykis tarp tėvinių klasių ir kitų poklasių. Siūlomos ontologijos struktūra atitinka konceptualiojo balanso principus.