An Ontology Model for Developing a SQL Personalized Intelligent Tutoring System

Wilairat Yathongchai[1,*], Jitimon Angskun[1] and Chun Che FUNG[2]

[1] School of Information Technology, Suranaree University of Technology, Thailand.

[2]School of Engineering and Information Technology, Murdoch University, Western Australia.

* Corresponding author. E-mail address: wilairat.bru@gmail.com

Received: 1 November 2016; Accepted: 26 January 2017

Abstract

SQL contents are important for learning and teaching in all disciplines of computer courses. This paper is a proposed ontology for the SQL-Personalized Intelligent Tutoring System (SQL-PITS). The SQL-PITS is an Intelligent Tutoring System (ITS) for teaching SQL and providing adaptive content according personalized information from the learners. It is considered as the best paradigm for tutoring learners with tutorial topics that consist of content, examples, exercises, and tutoring materials that are suitable for individual learners according to their abilities, profiles, preference and background. The SQL-PITS will present units of knowledge to be learned, which are called "Topics", as modular units separated from content and tutoring strategy. Ontology and Learning Object are used to enhance the capabilities of the SQL-PITS for effective SQL teaching. Ontology is a key concept in semantic web. It plays an important role in knowledge representation, sharing and reusing of multimedia learning materials, and content personalization. This paper presents the SQL ontology in three views: SQL Knowledge ontology, Learner Model ontology and Tutoring Strategies ontology. The structure of SQL ontology has been verified validity by domain experts which adopted the GQM (Goal, Questions, Metrics) approach for ontology evaluation. The evaluation results of SQL ontology structure in 5 ontology characteristics reveal that domain experts rate at the highest level on 4 ontology characteristics which are Preciseness, Completeness, Clarity, and Conciseness, The Consistency characteristic is in a high level.

Keywords: Ontology, Intelligent Tutoring System, Personalized learning, SQL

Introduction

The SQL is a special purposes programming language that is designed for managing data in the Relational Database Management System (RDBMS). It was constructed based on relational algebra and tuple relational calculus (Date, 2011). SQL contents are important for learning and teaching in all disciplines of computer courses. It is the most widely-used database querying language in the industry and it is expected that graduates should have mastered this skill (Kearns, Shead, & Fekete, 1997). SQL is a complete database language that contains data and view definition statements, as well as data manipulation statements. The SQL subjects need to be taught incrementally, with each concept builds on the previous one, from the specific to the general topics, in order to achieve a full understanding of the language and required concepts (Gennick, 2004). Despite its deceptively simple syntax, SQL is not a simple language to master. It is unlike the procedural or object-oriented languages, because SQL is essentially a declarative language, which allows us the developer to specify what is required and how to get it. SQL writers have to work with sets and reason about values. The concept of a group of values within a set not being ordered in any particular and learners must understand the concepts of values when they use it.

It is recognized that many learners have difficulties in learning SQL and this problem tends to persist throughout the course. This is partly due to the nature of SQL. There are also problems related to resources such as an increasing number of learners in a class, therefore the lecture may not be able to respond to all the queries. Other issues are related to the learners, such as lack of prior knowledge among the learners, their ability to learn, and their availability with respect to time. In addition, the SQL course have many topics such as SQL Syntax, SQL Operator, Select, Use of distinct and ordering, Grouping, Having, SQL function, Joining and so on, which have to be taught in sessions with fixed and limited time. Hence, it is important to find solutions to address the above issues. Renaud and van Biljon (2004a, 2004b) proposed two different pedagogical approaches to teach SQL and they compared the contrasts of the approaches in terms of mental models and cognition. In order to learn SQL more effectively, researchers (Soler, Boada, Prados, Poch, & Fabregat, 2009; Goldberg, 2009; Esendal & Dean, 2009; Piyayodilokchai, Panjaburee, Laosinchai, Ketpichainarong, & Ruenwongsa, 2013; Folland, 2016) have discussed the importance of theory and practice in the area of database and SQL learning.

SQL learning is essentially constructivist and the learners must identify the foundations of SQL that need to be built. The absence or insufficient understanding of these concepts will inevitably lead to a failure to comprehensively understand SQL. SQL-PITS provides knowledge by using adaptive tutorial, assessment, and learning materials that should be appropriate for the individual learners, to support the learner any time in the absence of human teachers, and to enable the learners to achieve better understanding, as well as complementing knowledge that is missing.

Ontology is an emerging instrument for knowledge representation, sharing, reuse and interoperability and it is playing an increasingly important role in the development of personalized intelligent tutoring architectures and systems. In this paper, the concept with respect to the ontology used as the knowledge base in an ITS is designed while the content development includes learning materials in the form of Learning Objects (LOs). The LOs are small instruction units that function as the learning materials in the ITS. Each unit of LOs has the self-contained and independent content which could be used and reused in various situations with other pieces of learning material and they can also be linked together to form new and larger size LOs. The LOs are applicable for organizing the structure of the sub-contents in accordance to the variety of learner's characteristics by ontology (Saidong, Guohua, Yaowen, & Yu, 2013). Ontology play a key role in deploying and reusing LOs on the Semantic Web. They are also used to describe metadata and other information related to the LOs with more flexibility and they facilitate the search and exchange of LOs. Ontology allow the specification of concepts in a domain and the terms used to markup content in LOs. The ontology shared allow for different systems to coming to a common understanding of the semantics of LOs (Mohan & Brooks, 2003). That is the main basis of knowledge base construction. These concepts are arranged in a hierarchical inheritance relationship and they have specific properties in each concept, which can be used to organize the structure of the LOs to support reusability and knowledge sharing (Shishehchi, Banihashem, & Zin, 2010) that can be used to collect and store the knowledge base in the Knowledge Repository. In addition, ontology has an ability to support smart search for selecting and allocating the appropriate LOs, and to build learning course according the learner model. ITS design is based on a variety of learning strategies and instructional design theory (Allen & Mugisa, 2010). The key roles of ontology in this paper, are the structural design with separate course content and LOs to enhance an ITS' ability to provide adaptive tutorials according to personalized learning of individual learner and reuse of LOs.

Methods and Materials

Architecture of an SQL Personalized Intelligent Tutoring System

The architecture of an SQL-PITS that comprises four modules: Knowledge Base module, Learner module, Pedagogical module, and Interface module. These modules are supported by ontology. It has been possible to promote a clear separation of the components of SQL-PITS, to make explicit the communication among the components, to indicate the support to build the personalized learning to the learners. Including with concept ontology, they make it possible to personalize learning to specific learners based on learners' knowledge level, preferences and learner' background. The architecture for an ontology-based SQL-PITS is shown in Figure 1 below.

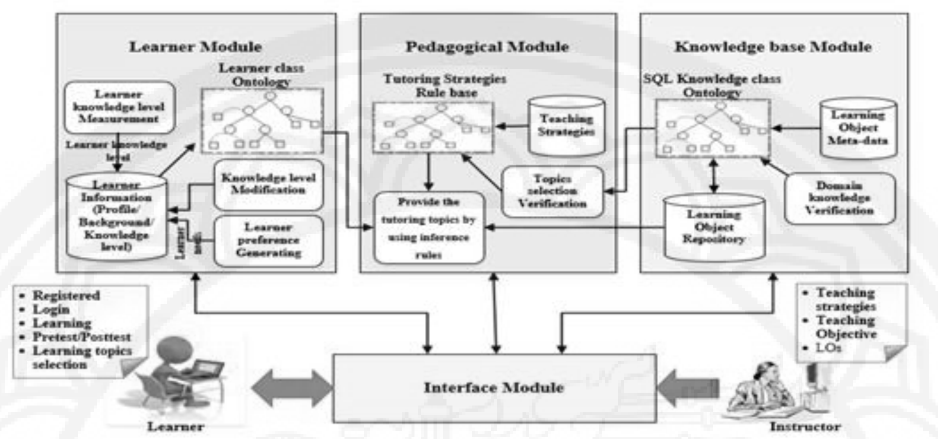

Figure 1 Architecture for ontology-based SQL-PITS

The Interface Module use to identify the intention of the learner and it is responsible for composing the information to be presented to the learner. It communicates with the Pedagogical Module with information about the learner such as profile, selected topics, learned topics and background. The Interface Module generates the graphic user interface for the lecturer to update the learning materials. The Learner Module contains information about the learner. This information is constantly updated and is used by the Pedagogical Module. The Pedagogical Module specifies the business rules of the system. Each strategy is a rule. The Pedagogical Module represents the learning objectives of a course and the tutoring strategies to be applied in each phase of the learning process. It interconnects every component in the system and it updates and consults the other components: Learner Module, Knowledge base Module, and Interface Module.

SQL Knowledge Ontology Model

The SQL knowledge ontology is the knowledge base of SQL contents for the SQL-PITS. The SQL knowledge domain is based on the computer science curriculum framework of the ACM/IEEE Computer Society (CM/IEEE-CS, 2013). The analysis of main concepts to design the SQL ontology is obtained from the learning outcomes (provided by the computer science curriculum framework), SQL tutoring objectives from the instructors (provided by the course of the academy), and the metadata of LOs (LOM) in accordance with the defined SQL knowledge domain. The analysis found that the unit of knowledge representation should provide in the *Topics* of tutorial which comprise Content Items, Example Items, Exercise Items, and Assessment Items. The SQL knowledge base, SQL concept map including SQL tutoring strategies, and SQL tutoring objectives were gathered from SQL instructors, books and websites. SQL experts have examined the validity of the materials by a Delphi technique (Judd, 1972).

The SQL knowledge ontology preserves the relationship between the Topics according to the SQL Concept map. A specific domain can have various interconnected Topics. A learner navigates the study from one Topic to another, through the relationship between them. Many kinds of relationships between Topics are possible. The structure of SQL ontology in view of SQL knowledge ontology is shown in Figure 2.

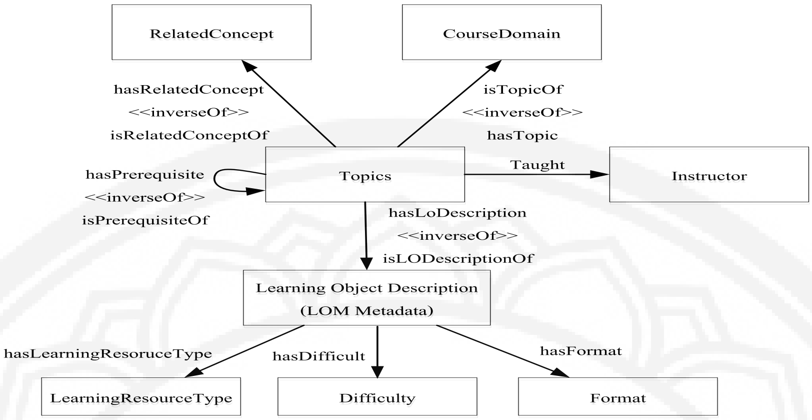

Figure 2 An excerpt of SQL Knowledge class ontology of SQL-PITS.

Learner Ontology Model

If a learner starts a tutorial session for the first time, the system will give him or her a set of pretest questions before the tutoring process starts. In this way, the Learner model estimates the new learner's background knowledge and determines the initial values. After the initial assessment of the learner, SQL-PITS recommends the relevant basic level of knowledge within the SQL knowledge domain and the learners can selected the *Topics* of tutoring by themselves. At the end of presented topic, the SQL-PITS gives the learner a posttest and assesses the learner's knowledge. During the session, the system adapts the learner's progress with the generated tutorial. If the learner answers the test items correctly, he progresses along the course and no change to the tutorial is necessary. However, if the learner fails to answer the test items correctly, the system will modify learner's performance. If the learner's performance does not meet expectations, the SQL-PITS will be dynamically re-planned. Through dynamic regeneration, each learner is able to get a highly personalized course for his needs.

The Learner Module contains information about the learner and progress of knowledge levels of the learner about the *Topics*. The learner needs to answer a psychological questionnaire which maps a set of 16 questions representing learning preferences and styles based on VARK tools (Fleming, 2016). The result indicates a preference of one of the four learning styles. These styles are stored in the Learner model. This is essential to adapt the material to the learner's characteristics. This information is constantly updated and is used by the Pedagogical Module and the Interface Module. The Learner ontology offers the opportunity to map all information about the learner. In this paper, four categories of learner's characteristics are considered which are 1) Learner Profile data - the learner's personal characteristics (name, ID, e-mail ...), 2) Learner ability data - the learner's cognitive and the weaknesses in the learners' knowledge, as well as individual characteristics, 3) Learner Preferences data - an individual's preferred way of learning using VARK model (Fleming & Mills, 1992) and 4) Background data - the current level of mastery of learners, the information log of tutoring system and attributes related to the corresponding elements in the Pedagogical Module. The structure of the Learner ontology is shown in Figure 3 below.

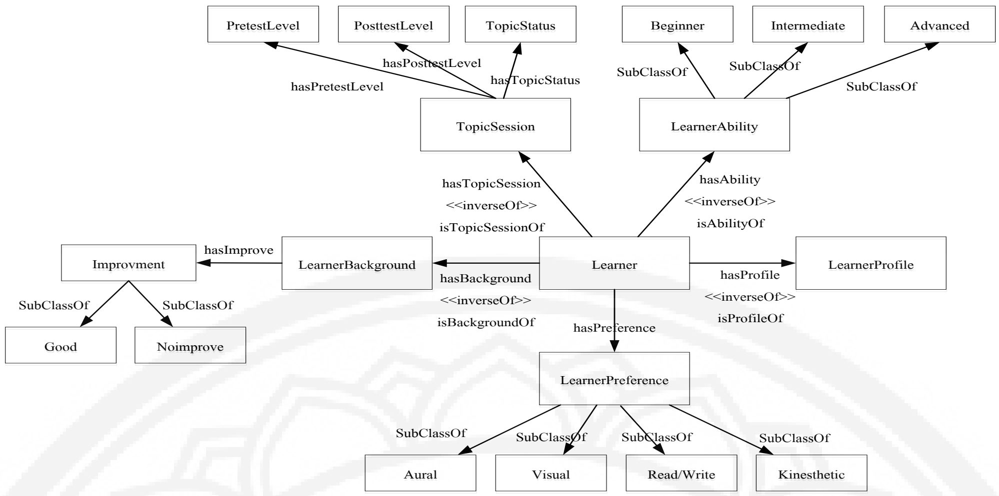

Figure 3 An excerpt of the Learner ontology of SQL-PITS.

SQL Tutoring Strategies Ontology Model

The common SQL pedagogical issues are to ensure that the fundamental concepts for the learner to acquire necessary background knowledge have been provided, and to give the learners an understanding of the syntax and semantics of SQL together with examples and exercises that are appropriate to the knowledge level of the learners. Methodologies for developing teaching expertise in AIED (Artificial Intelligence in Education) systems are based on 1) the observation of human teachers followed by an encoding of effective examples of these teacher's expertise, 2) learning theory, and 3) the observation of real students or on a runnable simulation model of the student (Du Boulay & Luckin, 2001).

In this paper, the tutoring strategy is determined based on gathered SQL teaching strategies using observation of a SQL teacher's teaching approach. The observations also included the learner's behaviors in the classrooms and interviews with teachers who have taught the SQL subject over three years. The collected data will be used to design the tutoring strategies using the Constructivism learning theory. The strategies could encourage the learners to achieve their learning objectives following the Bloom's Cognitive Domain (Bloom, 1976). SQL experts have also been invited to examine the validity of the tutoring strategies using the Delphi technique. The SQL-PITS tutoring strategies consist of two parts: *topic selection strategies* and *content presentation strategies* of the topics using examples, exercises, and tutoring materials. Determination of the conditions for selecting the appropriate strategies is considered from the Learner model. The ITS then checks the concept map to set the topic order of the prerequisite contents according to topics which the learner has not understood. The conditions to determine the strategies in each part are: 1) Evaluation results of the pretest levels of the learner in each topic which there are three levels: weak, medium, and good; 2) Evaluation results of the posttest levels in each topic that illustrates the progress of the learner's understanding of the SQL topic according to four levels: weak, medium, good, and no-result; and 3) A topic selection status which is either learned or not learned.

The Tutoring Strategies ontology determines the adaptation to be carried out based on a set of rules that associates the conditions and decisions. The conditions are composed of data from each of the components in the system.

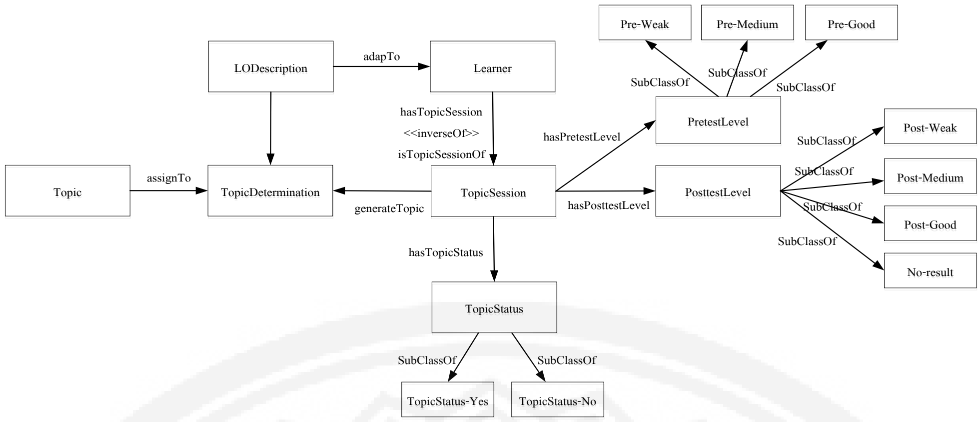

Figure 4 An excerpt of the Tutoring Strategies ontology of SQL-PITS.

SQL Tutoring Strategies Definition using SWRL Rules

This paper explicitly encodes SQL tutoring strategies as SWRL (Semantic Web Rule Language) rules. SWRL is based on a combination of the OWL DL and OWL Lite sublanguages of the OWL Web Ontology Language the Unary/Binary Datalog sublanguages of the Rule Markup Language. SWRL allows users to write Horn-like rules expressed in the term of OWL concepts to reason about OWL individuals (W3C, 2004). The SWRL rules can be used to infer new knowledge from the existing OWL knowledge bases.

Table 1 Examples of SQL Tutoring Strategies using SWRL rules

Predicate subclass	Pedagogical policy	SWRL Rule
Pre-Weak	a pretest score is less than or equal 49 and a period of pretest is within the limited time (PretestPeriod-In)	TopicSession(?x) ∧ PretestScore(?x, ?PreScore) ∧ swrlb:lessThanOrEqual(?PreScore, 49) ∧ PretestPeriod-In(?x) -> Pre-Weak(?x)
TopicDetermination	Learners who have preference in “Aural” group and their ability at an advanced level for each Topic then adapt SQL LOs to them in a Difficult level and Audio format.	Learner(?x) ∧ Aural(?x) ∧ hasTopicSession(?x, ?y) ∧ Advanced(?y) ∧ hasTopicOfSession(?a, ?y) ∧ hasLODescription(?a, ?z) ∧ hasFormat(?z, Audio) ∧ hasLearningResourceType(?z, Example) ∧ hasDifficult(?z, Very-difficult) -> adaptTo(?z, ?x) ∧ TopicDetermination(?z)

They are implemented as a SWRL rules that looks up each condition field. The advantage of modelling these policies as SWRL rules is that each policy’s computation is explicitly represented in the ontology, and can be viewed and edited, as well as reasoned about by other applications.

Results and Discussion

This paper adopted the GQM (Goal, Questions, Metrics) approach for ontology evaluation. This approach, using the metrics is derived from goals with the formulated questions that allow measuring to concentrate on requirements for SQL ontology rather than on data (Basili, Caldiera, & Rombach, 1994). The processes have 3 steps as:

Step 1: Identification of goals for SQL ontology. The proposed ontology will provide the knowledge to support individual learning outside the classrooms, by considering the characteristics of the learners in the term of their prior knowledge on SQL, preference, and their learning ability. These characteristics deal with the ontological background are completeness, consistency, conciseness, preciseness and clarity (Tankelevicienea & Damasevicius, 2009).

Step 2: Formulation of questions based on the definitions and descriptions of desired characteristics. After the ontology characteristics have been identified, the questions are formulated, which an expert will perform their evaluation. Achievement of every desired goal is checked by several questions, where every question is matched by one evaluation value. The values of metrics from one group are used to calculate a measure of one desired characteristic.

Step 3: Definition of metrics, which are based on the human evaluation. The SQL ontology structure evaluation should be used the 5 experts of SQL and ontology, because they can capture the semantic meaning of ontology and evaluate its quality in the context of SQL domain knowledge. The measurement system for the SQL ontology characteristics uses a five point Likert scale (Likert, 1967). The evaluation value of a characteristic based on a group of questions is calculated as an average, Standard deviation (S.D.), and expressiveness level of evaluation results for that group.

The presented expert evaluation form based on ontology quality criteria provides a set of questions guiding the ontology evaluation process and combining the results of evaluation by experts.

Table 2 Evaluation results of SQL ontology quality by experts.

Characteristics	SQL ontology quality evaluation metric		
	Mean	S.D.	Expressiveness Level
Completeness	4.40	0.53	Strongly agree
Consistency	4.20	0.45	Agree
Conciseness	4.30	0.63	Strongly agree
Preciseness	4.47	0.50	Strongly agree
Clarity	4.40	0.55	Strongly agree

The evaluation results of the ontology structure show a good classification, a structural completeness, and an internal relation of the ontology structure. This result leads to organizing the tutorial topic selection including the LOs presentations as each learner's characteristics.

Conclusion and Suggestion

In this paper, an architectural design of the SQL-PITS that uses ontology for the building of a SQL tutoring system is proposed. The SQL-PITS is a framework for an adaptive and personalized tutoring system. The formats of several ontologies have been proposed and they correspond to the components of a tutoring system. The explicit conceptualization of the system components is in the forms of ontology and it facilitates knowledge sharing, knowledge reuse, communication, and construction of rich knowledge and intensive systems. Ontology and Learning objects are used for the construction of the SQL-PITS. They can interoperate and integrate to provide the learners with adaptive contents.

The SQL-PITS can provide tutorial topics that consist of contents, examples, exercises, and tutoring materials. The tutorial topics are suitable for individual learners according to their abilities, profiles, preference, and background. The SQL-PITS is used to supplement learning outside the classrooms according to a learner's personal learning individually. It also focuses on the provision of consistently relevant and interrelated contents to the learners. The tutorial topics are associated with a concept map to support the learners by offering tutorials that address the weaknesses in the learners' knowledge. The system can recommend all relevant basic level of knowledge within the SQL knowledge domain.

All materials for design and development of an SQL ontology such as SQL knowledge base, SQL Concept map, SQL tutoring strategies and SQL tutoring objectives were gathered from the SQL instructors, books, and websites. The validity of the materials has been examined by SQL experts with the Delphi technique. The evaluation results of SQL ontology structure in 5 ontology characteristics reveal that the ontology domain experts rate at the high level which adopted the GQM approach for ontology structure evaluation.

References

Allen, C. A., & Mugisa, E. K. (2010). Improving Learning Object Reuse Through OOD: A Theory of Learning Objects. *Journal of Object Technology*, *9*(6), 51-75.

Basili, V., Caldiera, G., & Rombach, H. D. (1994). Goal question metric approach. In J. J. Marciniak (Ed.), *Encyclopedia of Software Engineering* (pp. 528-532). New York: John Wiley and Sons.

Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGrawHill.

CM/IEEE-CS Joint Task Force for Computer Curricula 2013. (2013). Computing Curricula 192 2005: An Overview Report. Retrieved from http://www.acm.org/education/curric_vols/Cs2013-ComputerScience Curricula 2013.pdf.

Date, C. J. (2011). *SQL and Relational Theory: How to Write Accurate SQL Code*. USA: O'Reilly Media, Inc.

Du Boulay, B., & Luckin, R. (2001). Modelling human teaching tactics and strategies for tutoring systems. *International Journal of Artificial Intelligence in Education*, *12*(3), 235-256.

Esendal, T., & Dean, M. (2009). A Model to Make the Learning of Internet Programming with Databases Easy. In *7th Workshop on Teaching, Learning and Assessment in Databases, Birmingham, UK, HEA*.

Fleming, N. D. (2016). The VARK questionnaire 7.1. Retrieved from http:// www.vark-learn.com/

Fleming, N. D., & Mills, C. (1992). Not Another Inventory, Rather a Catalyst for reflection. *To Improve the Academy*, *11*, 137-55.

Folland, K. A. T. (2016). *viSQLizer: An interactive visualizer for learning SQL*. (Master's thesis). Norwegian University of Science and Technology (NTNU), Norway.

Gennick, J. (2004). On the importance of mental model. Retrieved from http://gennick.com/mental models.html.

Goldberg, C. (2009). Do you know SQL? About Semantic Errors in Database Queries. In *7th Workshop on Teaching, Learning and Assessment in Databases, Birmingham, UK, HEA*.

Judd, R. C. (1972). Use of Delphi methods in higher education. *Technological Forecasting and Social Change*, *4*(2), 173-186.

Kearns, R., Shead, S., & Fekete, A. (1997). *A teaching system for SQL: Proceedings of the 2nd Australasian conference on Computer Science education, 2-4 July 1997* (pp. 224-231). Melbourne, Australia: University of Sydney.

Likert, R. (1967). The Method of Constructing and Attitude Scale. In Reading in Fishbein. M.Ed., *Attitude Theory and Measurement* (pp. 90-95), New York: John Wiley & Sons.

Mohan, P., & Brooks, C. (2003). *Learning Objects on the Semantic Web: Proceedings of The 3rd IEEE International Conference on Advanced Learning Technologies (ICALT 2003), 9-11 July 2003* (pp. 195-199). Athens, Greece: IEEE.

Piyayodilokchai, H., Panjaburee, P., Laosinchai, P., Ketpichainarong, W., & Ruenwongsa, P. (2013). A 5E learning cycle approach-based, multimedia-supplemented instructional unit for structured query language. *Journal of Educational Technology & Society*, *16*(4), 146-159.

Renaud, K. V., & Van Biljon, J. A. (2004a). *Teaching SQL — Which Pedagogical Horse for This Course?: Proceedings of a 21st British National Conference on Databases, BNCOD 21, 7-9 July 2004* (pp. 244-256). Edinburgh, UK.

Renaud, K., & Van Biljon, J. (2004b). Teaching SQL—Which Pedagogical Horse for This Course? In *British National Conference on Databases* (pp. 244-256). Springer Berlin Heidelberg, Germany.

Saidong, L., Guohua, T., Yaowen, X., & Yu, S. (2013). The Research and Implementation of a Personalized Intelligent Tutoring System: Proceedings of *the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)* (pp. 1302-1304). France: Atlantis Press

Shishehchi, S., Banihashem, S. Y., & Zin, N. A. M. (2010). *A proposed semantic recommendation system for e-learning: A rule and ontology based e-learning recommendation system. Proceedings of the 2010 International Symposium in Information Technology (ITSim), 15-17 June 2010* (pp. 1-5). Kuala Lumpur: IEEE.

Soler, J., Boada, I., Prados, F., Poch, J., & Fabregat, R. (2009). A web-based e-learning tool for the continuous assessment and skills improvement of main database course topics. In *7th Workshop on Teaching, Learning and Assessment in Databases, Birmingham, UK, HEA*.

Tankelevicienea, L., & Damaseviciusb, R. (2009). Characteristics of Domain Ontologies for Web Based Learning and Their Application for Quality Evaluation. *Informatics in Education*, *8*(1), 131-152.

W3C. (2004). SWRL: A Semantic Web Rule Language Combining OWL and RuleML. Retrieved from https://www.w3.org/Submission/SWRL/