DBLearn: Adaptive E-Learning for Practical Database Course – An Integrated Architecture Approach

Srinual Nalintippayawong, Kanokwan Atchariyachanvanich, Thanakrit Julavanich
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
srinual@it.kmitl.ac.th, kanokwan@it.kmitl.ac.th, 56070053@kmitl.ac.th

***Abstract*— In this paper, an integrated architecture approach in designing and developing a DBLearn web-based application is presented. The DBLearn system is a personalized and adaptive e-learning system designed especially for learning practices in database courses. This approach focused on topics that are important but difficult for new learners, such as database design and structured query language (SQL) command query. The concept of adaptive e-learning and autonomous agents were applied in this system to eliminate the traditional constraints of effective e-learning, such as the problem of different learning sensory and knowledge levels. Four approaches were used to solve this problem. First, learning style theory was used to classify the way of learning for each student. Second, the student activity (historical data) is kept in the system to analyze the next knowledge the student should learn or review. Next, the SQL query automated grader was used to judge the correctness of the student's query. This grader supports all the necessary commands in both DML and DDL. Finally, the SQL query question generator module that can generate SQL query questions automatically is presented. This will reduce the instructor's work load in creating enough questions and allow the students to practice at their own pace as much as they want. By using these four techniques, the students will have a better learning experience and becoming more successful in learning outcomes.**

Keywords—e-learning; personalized learning; adaptive learning; database design; structured query language; learning outcome

I. Introduction

Information technology has become one of the most significant parts of our lives, no matter what form it is in, such as devices, software, websites, or others. The most important part that has driven these technologies is the data stored in the database. For example, a social media website stores member's data (profiles, posts and photos) in the database and connects to the programming language, like PHP and JSP, to display the information on the page. As another example, convenience stores keep track of their stock, customer receipts, etc., to improve their stock management and help with decision-making, such as ordering new stocks. Therefore, the knowledge in the database field, like relational database and structured query language (SQL), becomes very important for a career related to technology, such as a programmer, data analyst, computer science researcher, and so on.

In the era of electronic devices and the internet, learning computer language from an e-learning system is very popular nowadays due to the easiness of access. Many e-learning system websites help their students by embedding the integrated development environment (IDE) on their websites. The embedded IDE helps students to learn and try to code instantly on the website. It also can check the correctness of the answers given by the student and give advice back to improve or correct the answer. But the problem is some computer languages, like SQL, still lack fully functional embedded IDEs that can check the correctness of an answer. There are some integrated IDEs that have been developed for SQL that students can use some commands with, but they do not cover all the SQL commands that students should know [1, 2]. Moreover, to learn all the essential relational database knowledge, students need to learn how to design a database correctly, but while this is very important it is difficult for new learners [3, 4]. Moreover, it is hard to find a suitable tool to integrate with the e-learning system. These problems make the learning of database knowledge via an e-learning system not fully effective.

II. Research Problems

There is a difference in each learner, with the two most important differences being the learning approach and knowledge level. Accordingly, instructors need to customize the content to suit each learner [5]. These are the barriers in studying via an e-learning system that means it will not fully fit every learner, and so they have became the root of personalized and adaptive e-learning research. Personalized e-learning customizes the learning experience to be most suitable for each student. However, there are still some problems in personalized e-learning. One of the largest problems is that the instructor has a huge work load to create the learning objects, both contents and questions, to feed the personalized algorithm to reach its full potential. Autonomous agents can effectively manage the learning process of each learner [5].

978-1-5090-5504-3/17/$31.00 ©2017 IEEE
SNPD 2017, June 26-28, 2017, Kanazawa, Japan

To overcome this, DBLearn was developed. This is an integrated architecture and approach in a web-based system for adaptive e-learning in a practical database course. The system supports practices in database development topics, such as database design and SQL command query. It uses learning style theory, SQL query grading system, database design tool and SQL query question generation algorithm in conjunction to help learning and teaching via an adaptive e-learning system that hence becomes more efficient and effective.

III. Related Works

A. E-Learning Personalization

E-learning personalization is the method that creates a different learning experience for each student. The experience is created by a specific algorithm to adapt each student's learning experience. There has been a lot of research in the e-learning personalization field, and these are summarized for the personalization approaches with examples of its implementation as follows.

a) Learning style adaptation: In 2013, Bayasut et al. [1] used learning style theory to personalize the experience for each student. The learning style theory aims to account for the differences in each student's learning progress, where every student can be classified into one or more of the proposed styles of learning. In this research, the system uses twelve variables from the student's e-learning browsing (historical) data to classify them. The system will present the best learning content, determined by a number of style-matched objects in the content, to match the result.

b) Learning path adaptation: In 2015, Alshalabi et al. [2] developed an adapted learning path technique to personalize students learning. The system models knowledge domains in the course to a connected graph structure and uses a shortest path algorithm to find a way to present the content to help students reach their goal as fast as possible. According to this approach, students will have their own learning experience depending on each student's goal. In addition, in 2009, Brusilovsky et al. [8] used adaptive navigation support to guide students to the appropriate learning object by using a partially filled bullet icon to denote the level of user knowledge for each topic.

c) Assessment adaptation: In 2013, Jadhav et al. [9] proposed the assessment generation and question adaptation technique to personalize students' learning according to their knowledge level and performance using their historical data. The technique helped students to improve their knowledge and understanding of each knowledge domain better.

B. Personalized E-Learning Architecture

From the previous e-learning research, there are three important modules [6, 10], which are described as follows.

a) Domain module: Store and manage learning objects, such as the contents and questions, available in the system. In addition, it assists the instructor to create learning objects, design the knowledge model and evaluate a suitable learning style for each content.

b) User module: Store and manage the user profile and all user-related data, such as personal profile, learning style and activity log. The module will use the data to create a user model for use in conjunction, as both input and output, with the adaptation module.

c) Adaptation module: Analyze the best learning object for the student by using data from the domain and user models together. The module will present the learning object to the student via the interface.

C. E-learning System for Database Course

From the previous research [1, 11, 12], the e-learning systems that are designed specifically for the database course are likely to have an automatic SQL query grader implemented in the system. But the graders support only a few commands like SELECT. Some other necessary commands are not supported yet due to the difficulties in correctness validation or security issues.

Another important content in database course learning and teaching is the database design part, such as designing an Entity-Relationship (ER) model and transforming the ER model into the relational model. There is still no current system that implements any tool to help students and instructors in this part. By implementing this, the learning and teaching database course via e-learning will become more complete. The feature comparison between traditional systems and the proposed system of this study is shown in Table 1.

TABLE I. Feature Comparison Between Traditional Systems and The Proposed System of This Study

Feature	**XDa-TA [1]**	**SQLify [11]**	**ULUL-ILM [6]**	**SQLator [12]**	**ADAPT^2 [7]**	**Proposed system**
Adaptation	No	No	Yes	No	Yes	Yes
Learning style	-	-	Yes	-	No	Yes
Learning guidance	-	-	No	-	Yes	Yes
Database course tools	Yes	Yes	No	Yes	Yes	Yes
SQL grader	Yes	Yes[a]	-	Yes[a]	Yes[a]	Yes[b]
Database design tool	No	No	-	No	No	Yes
SQL query question generator	No	No	-	No	Yes[a]	Yes[b]

[a] Support only SELECT commands.

[b] Support all data manipulation language commands and necessary data definition language commands.

D. VARK Learning Style

Learning style theory aims to classify the learning characteristic of each person. For example, some students like to learn a new area of knowledge by doing their practice by themselves in real life more than just reading theory from books, while other students may prefer an explanation by photos or diagrams more than by descriptive text, or visa versa. Since the learning style became a research topic, there have been several diverse learning style theories proposed, such as such as the Felder-Silverman [13], Kolb [14], Myers-Briggs Type Indicator [15] and VARK [16] learning styles.

The Vark learning style theory [16], which was used in this system, categorizes students into one of four learning styles. First, *visual learners* are students who perceive and learn by visualization using images, diagrams or flow charts. Second, *aural learners* are students who perceive and learn by hearing, listening or communicating with others. This type of learner likes to learn by listening to lectures from instructors or recorded sessions. Third, *read/write learners* are students who perceive and learn by writing and reading books or articles. Most of time, they are likely to take notes for reviewing later. Lastly, *kinesthetic learners* are students who perceive and learn by examples, either simulated or real-life examples. The VARK learning style theory is multimodal, where each learner can be classified into one or more of these learning styles.

The VARK learning style theory will classify students using a questionnaire comprised of sixteen four-choice questions. Students can answer more than one choice in each question. Four choices represent the way to approach or think for each learning style in the situation addressed in the question. The evaluation of the questionnaire can be done by counting each answer as the preference score.

The advantage of the VARK learning style theory is it classifies each student according to their learning sensory, which makes each style clearly different from the others. There is also the advantage for the instructor in that, since there are only four learning styles, the instructor will not have much difficulty or work to create the learning objects to suit each style. Moreover, the classification uses the least amount of questions in the questionnaire, which prevents the student from losing interest and giving inaccurate answers from inattention.

IV. THE ARCHITECTURE OF THE SYSTEM

There are five important modules in the DBLearn system, as shown in Figure 1. Each of them takes care of a different part of the system as described below.

A. Learner Profile Module (LPM)

The LPM stores the profile data of all the users, such as the name, enrolled course, learning style preference (obtained from the result of learning style questionnaire) and knowledge data (obtained from the historical usage in the system).

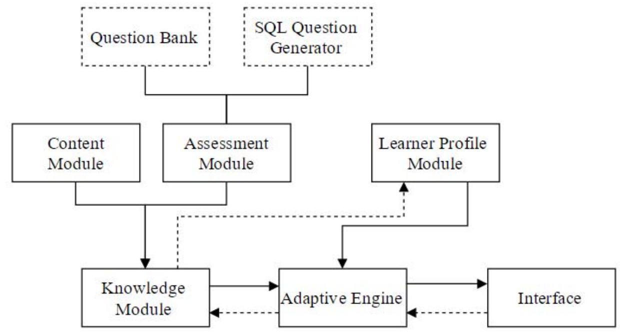

Fig. 1. The architecture of DBLearn system

B. Content Module (CM)

The CM stores all the contents available in the system, and categorizes the contents by knowledge domain and course. The content stored in the CM has tags to determine the knowledge domains associated with this content and learning style that is suited for this content.

C. Assessment Module (AM)

The AM stores all the questions available in the system. Each question also has knowledge domain tags like the content in the CM. The AM categorizes questions using the knowledge domain. Thus, in addition to handpicked questions in the question set from the instructor, the system can create a question set by picking questions adaptively to match each student's knowledge and performance. The AM can also generate SQL query questions automatically.

D. Knowledge Module (KM)

The KM gathers the contents and questions from the CM and AM and categorizes the contents and questions into each KM that has been tagged in each content or question. The KM will communicate with and support Adaptive Engine (AE) and LPM to present the best material for the learner.

E. Adaptive Engine (AE)

The AE will communicate with the interface, KM and the LPM. The AE retrieves a list of knowledge from the KM and displays them through the interface. When the learner selects the knowledge that they wanted to learn, the AE will display suitable content according to the learning style in the LPM. The AE also retrieves recorded activity from the LPM to analyze which knowledge the learners should study or review next, either in the form of contents or questions.

V. APPROACH AND DEVELOPMENT

The approach and development for each module in the DBLearn system can be described as follows.

A. Knowledge Domain

The DBLearn system stores all the knowledge of the course. Each knowledge needs to be identified for its prerequisite knowledge, which will make the system available to map into the connected graph model. For example, in Figure 2, the knowledge domain model was mapped from the curriculum in the first course to test DBLearn system, in this case the database system concepts course of the Faculty of Information Technology, KMITL.

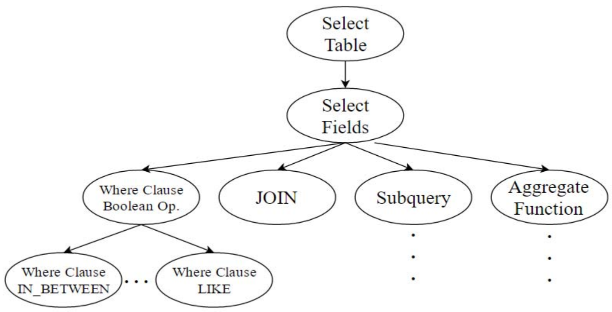

Fig. 2. Example of knowledge domain model

B. Contents

There are four types of knowledge contents in the DBLearn system (text, text with an example, video and audio). Each type of content has been designed according to each learning style of VARK theory. For the content delivery, a content type priority technique is used where a priority level is assigned in each content type for each learning style, to deliver the available content with the highest priority to a student. The priority for each content type is shown in Table 2.

In this system, one knowledge domain can contain one or more contents, and there are four content types for each content as described above. The knowledge and content structure is summarized in Figure 3.

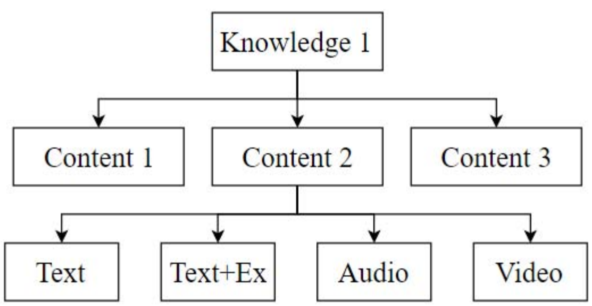

Fig. 3. The knowledge and content structure of the DBLearn system

C. Assessments

In this system, six types of questions (multiple choice, fill in the blank, matching, SQL query, ER design and ER transformation) were developed. These six types of assessment will help students to achieve the learning goals of a database course. The types of question are described as follows.

a) Multiple choice questions: Students have to answer one or more choice correctly to get the point.

TABLE II. PRIORITY OF CONTENT TYPE PRESENTATION FOR EACH LEARNING STYLE

Learning style	Priority of content type			
	Text	***Text + Ex.***	***Video***	***Audio***
Visual	Low	Low	High	Medium
Auditory	Low	Low	Medium	High
Read/Write	High	Medium	Low	Low
Kinesthetic	Medium	High	Low	Low

a) Fill in the blank: Students have to answer by typing in the input field correctly to get the point. Instructors may choose to score the answer later themselves, if there can be more than one possible answer or it is an open-ended question.

b) Matching: Students have to match the keywords given that have a relationship with each other. There may be more than one relationship for one keyword, but students have to answer just one. The system will check the correctness automatically. Instructors also can make the students explain the relationship in the answer, but then the instructors have to check the correctness of the explanation themselves later.

c) SQL Query: Students have to write a SQL query to get the result as specified in the question. The system will check the correctness automatically from the result returned after executing the queries, not from the query itself, because there are many ways to write a query to get the expected result. The system supports all data manipulation language (DML) commands and some of the important data definition language (DDL) commands, such as CREATE, ALTER and DROP.

d) ER Model Design: Students have to design the Relational Database using an ER model from defined business rules. Students can use the toolbox embedded in the system for modeling. For the correctness checking, if a student's answer is different from the correct answer, the system will show the difference to instruct for scoring. The system will not score the answer automatically because of the variety of possible correct answers. This problem can be solved in the future.

e) ER Transformation: For the design process, the student has to transform the ER/EER model into the relational schema correctly, which includes multiple relations, attributes, primary keys and foreign keys. Students have to answer accurately all of these schema elements to get the full point. For the scoring method, a rule-based scoring, as described in Table 3, was used. Instructors can define the score for each schema element. The system will show the difference from the correct answer to the student.

To help solve one of the stated problems, the difficulties in creating enough questions to suit all the student levels of knowledge, the DBLearn system helps instructors by generating SQL query questions from the database template. The system creates a correct query from SQL commands and difficulty level inputs. After that, the system creates a question description of each part of the query using the description pattern and display to the student.

TABLE III. EXAMPLE FOR ER TRANSFORMATION ANSWER SCORING

Element	Score		Threshold[a]
	Full score	***Incorrect***	
Relation	0.00	-0.50	-
If Relation name is correct:			
Attribute	0.50 total	-0.15 each	70%
Primary Key	0.50 total	-0.50 each	100%
Foreign Key	0.50 each	-0.50 each	100%

[a.] Student must answer that element correctly enough to reach the threshold to get the point.

D. Personalization

The personalization will be processed by the AE using three techniques. First, the learning styles, which were classified using the results of the 16-multiple choice questionnaire. When a student chooses the knowledge that they want to learn, the system will deliver the best content according to the student's learning style, as shown in Table 2. If a student is not satisfied with the delivered content, the student can change the style of the content himself. Students can also retake the learning style classification questionnaire to change their learning style permanently.

The second technique is the knowledge guidance. The system keeps track of the activities while the student browses through the courseware or answers the questions. If the student answered questions in the same knowledge domain incorrectly above a certain threshold level in that session, the system will suggest that the student should review this knowledge domain again. When the correct answer rate is satisfied, the system will suggest the best knowledge domain to learn next using data from the knowledge domains' path in the KM.

Finally, the third technique is the personalized SQL query assessment. The system generates the assessment as a set of questions according to the student's knowledge model stored in the LPM. The generated assessment will improve and review the understanding level in each learned, but not yet understood, knowledge domain first. It then reviews the learned and understood knowledge domain. If that knowledge domain is about the SQL query the system will use the SQL query question generator to create question from the database schema in the system, to make a practicable environment and avoid duplicate questions.

VI. SYSTEM EVALUATION

The DBLearn system is currently in development. The SQL query assessments module with knowledge prediction was chosen for evaluation. The evaluation was conducted to assess how accuracy of the prediction was, after the SQL query test was done with 121 sophomore students in Faculty of Information Technology, KMITL. The module is to predict whether each student understands each SQL query in the knowledge domain or not. The students have studied basic database concepts from the theoretical classes and SQL querying from practical classes. Each student was asked to answer twenty-one SQL query questions. The screenshot of SQL query test is shown in Figure 4. We gathered answers and converted them to be information in an unlearned knowledge domain of each student. The next time the students do the exercise of SQL query, this information will be used to suggest questions that suit for each student. After finishing the test session, each student was asked to answer an online questionnaire for assessing whether they understood each knowledge domain or not. The results from the SQL query assessments module with knowledge prediction and from online questionnaires were analyzed using a confusion matrix. The summary result of system performance is shown in Table 4.

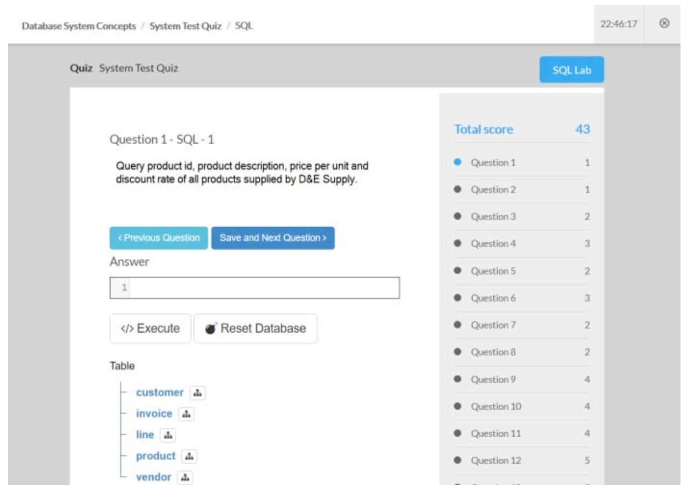

Fig. 4. Example screenshot of the system

TABLE IV. PERFORMANCE OF KNOWLEDGE PREDICTION USED IN ADAPTIVE MODULE

Knowledge domain	Precision	Recall	F-Measure
SELECT	100.00%	91.49%	95.56%
JOIN	100.00%	74.12%	85.14%
WHERE	97.01%	79.27%	87.25%
ORDER BY	98.84%	88.54%	93.41%
Aggregate function	97.17%	98.10%	97.63%
GROUP BY	98.48%	78.31%	87.25%
Subquery	98.28%	71.25%	82.61%
Total	**98.17%**	**83.84%**	**90.44%**

The results in Table 4 revealed that the prediction algorithm used in the adaptive module produced satisfactory results because the F-Measure of each knowledge domain is greater than 80%. There are some knowledge domains, for example, JOIN and subquery, having lower than 80% of recall. It maybe because the questions were complicated and required combined knowledge domains. For instance, the students probably already understood JOIN but answered incorrectly because of other commands. Therefore, the system returned wrong predictions. In the future, we will improve the prediction algorithm by finding actual mistakes in student's answer and preparing more precise questions to increase the accuracy of the prediction.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, an integrated architecture and approach was proposed in the design and development of DBLearn, a web-based adaptive e-learning system designed especially for database and SQL related courses. In the system, the student's learning style is used to help present content adaptation. In addition, the student activity (historical data) is used to create the content suggestion, while the SQL query question generator is used to generate SQL query questions automatically. This will allow students to keep practicing with suitable questions as much as they want. The summary assessment module has a variety of practical types, as shown in Figure 5.

The DBLearn system provides several advantages over current systems, such as (i) solving the major problem of teaching a database course to students of different learning styles and knowledge levels, and so resulting in suitable contents and achievable learning outcome; (ii) helping students to practice the SQL query and database design more efficiently and effectively using the provided web-based tools; (iii) and the autonomy grading and error explanation. In addition, the difficulty for instructors to create enough questions is eliminated since the students can let the system automatically suggest questions that they should keep practicing on. This not only allows students to have a better learning experience and become more successful in learning the database course, but the instructors will have less work in adaptive e-learning management.

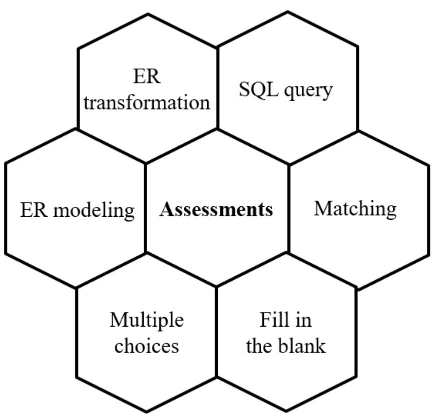

Fig. 5. Types of assessment in DBLearn

In the future, the DBLearn system design and development will focus on the design and development of the automatic SQL query question generator algorithm, and develop the algorithm to solve the problem of autonomous grading of the ER modeling. A final evaluation of the developed system to show that it really adapts to suit each student and provides them have a better learning experience and improved knowledge is required. One further improvement to this system is to allow for students' typing errors to reduce the workload for manual grading. students' typing errors to reduce the workload for manual grading. Finally, find out reinforcement learning-based algorithms to motivate the practice intensity as high as in the classroom environment [17].

ACKNOWLEDGMENTS

Research presented in this paper was supported by the Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang (KMITL) [grant number 2558-06-008]. The authors would like to thank the Research Funding Board of the Faculty of Information Technology, KMITL.

REFERENCES

[1] A. Bhangdiya, B. Chandra, B. Kar, B. Radhakrishnan, K. V. M. Reddy, S. Shah, and S. Sudarshan, "The XDa-TA system for automated grading of SQL query assignments." pp. 1468-1471.

[2] C. Kleiner, C. Tebbe, and F. Heine, "Automated grading and tutoring of SQL statements to improve student learning," in Proceedings of the 13th Koli Calling International Conference on Computing Education Research, Koli, Finland, 2013, pp. 161-168.

[3] C. J. Date, Hugh Darwen, and Nikos A. Lorentzos "Database Design I: Structure Time and Relational Theory", Second Edition, 2014, pp. 197-226.

[4] C. Domínguez and A. Jaime. "Database design learning: A project-based approach organized through a course management system", Original Research Article Computers & Education, Volume 55, Issue 3, November 2010, pp. 1312-1320.

[5] D. Benyon and D. Murray. Adaptive systems: from intelligent tutoring to autonomous agents Original Research Article Knowledge-Based Systems, Volume 6, Issue 4, December 1993, Pages 197-219.

[6] B. L. Bayasut, G. Pramudya, and H. B. Basiron, "ULUL-ILM: The design of web-based adaptive educational hypermedia system based on learning style." pp. 147-152.

[7] I. A. Alshalabi, S. Hamada, and K. Elleithy, "Automated adaptive learning using smart shortest path algorithm for course units." pp. 1-5.

[8] P. Brusilovsky, A. Mitrovic, S. Sosnovsky, M. Mathews, M. Yudelson, D. H. Lee, and V. Zadorozhny, "Database exploratorium: a semantically integrated adaptive educational system," Ubiquitous User Modelling, pp. 24, 2009.

[9] M. Jadhav, S. Rizwan, and A. Nehete, "User profiling based Adaptive Test Generation and Assessment in E - Learning System." pp. 1425-1430.

[10] S. Hamada, I. A. Alshalabi, K. Elleithy, and I. A. Badara, "Automated Adaptive Mobile Learning System using the Semantic Web." pp. 1-7.

[11] M. d. Raadt, S. Dekeyser, and T. Y. Lee, "Do students SQLify? improving learning outcomes with peer review and enhanced computer assisted assessment of querying skills," in Proceedings of the 6th Baltic Sea conference on Computing education research: Koli Calling 2006, Uppsala, Sweden, 2006, pp. 101-108.

[12] S. Sadiq, M. Orlowska, W. Sadiq, and J. Lin, "SQLator: an online SQL learning workbench," in Proceedings of the 9th annual SIGCSE conference on Innovation and technology in computer science education, Leeds, United Kingdom, 2004, pp. 223-227.

[13] R. M. Felder, and L. K. Silverman, "Learning and teaching styles in engineering education. Engineering Education," 1988.

[14] A. Y. Kolb, and D. A. Kolb, The Learning Style Inventory: Technical Manual, Boston, MA: McBer & Co, 1976.

[15] "MBTI® Basics," http://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/.

[16] "The VARK Modalities," http://vark-learn.com/introduction-to-vark/the-vark-modalities/.

[17] C. van Klaveren, S. Vonk, and I. Cornelisz, "The Effect of Adaptive versus Static Practicing on Student Learning-Evidence from a Randomized Field Experiment," Economics of Education Review, 2017.