

กลไกการสร้างแบบทดสอบอัตโนมัติสำหรับคำถามแบบปรนัย

โดยการประยุกต์ใช้ข้อมูลแบบออนโทโลยี

Automatic Quiz Generation Mechanism for Multiple Choices Question

by Applying Ontological Data

ศศิธร หนูทอง^{1*}, สุนทร วิฑูรพจน์², เบญจพร หนูทอง³

Sasitorn Nuthong^{1*}, Suntorn Witosurapot², Benjaporn Nuthong³

¹วิทยาลัยการคอมพิวเตอร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต

²ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์

³ศูนย์วิทยาศาสตร์การแพทย์ที่ 11/1 ภูเก็ต กรมวิทยาศาสตร์การแพทย์

¹College of Computing, Prince of Songkla University, Phuket Campus, Thailand

²Department of Computer Engineering, Faculty of Engineering, Prince of Songkla University, Thailand

³Regional Medical Sciences Center 11/1 Phuket, Department of Medical Sciences, Thailand

*Corresponding author. Tel.: 08 4750 2298, E-mail: sasitorn.n@phuket.psu.ac.th

บทคัดย่อ

ระบบการสร้างแบบทดสอบเป็นระบบที่มีความสำคัญมากในด้านการศึกษา เนื่องจากช่วยในการประเมินความรู้ความเข้าใจในบทเรียนของนักเรียน รวมทั้งยังช่วยให้ผู้ใช้งานสามารถสร้างแบบทดสอบได้อย่างสะดวกและรวดเร็ว แต่ระบบการสร้างแบบทดสอบที่มีอยู่ในปัจจุบัน ยังมีข้อจำกัดเรื่องความสามารถในการสร้างข้อคำถาม ซึ่งยังเป็นการป้อนข้อมูลเข้าสู่ระบบด้วยผู้ใช้งานเองทั้งหมด ทำให้ต้องสิ้นเปลืองเวลาในการสร้างแบบทดสอบเป็นเวลานาน งานวิจัยนี้จึงมีวัตถุประสงค์ในการนำเสนอกลไกในการสร้างแบบทดสอบอัตโนมัติสำหรับคำถามแบบปรนัย โดยการประยุกต์ใช้ข้อมูลแบบออนโทโลยี เพื่อนำมาใช้ในการประเมินระดับความยากง่ายของคำถาม โดยใช้วิธีการวัดค่าความคล้ายคลึงแบบลูกผสมด้วยการผสมระหว่างเทคนิควิธีการวัดค่าความคล้ายคลึงเชิงความหมาย ความสัมพันธ์เชิงความหมาย และค่าคุณสมบัติของข้อมูล เพื่อนำมาคำนวณหาค่าระดับความยากง่ายของคำถาม จากนั้นจึงนำกลไกที่นำเสนอมาประยุกต์ใช้งานโดยใช้เครื่องมือ RDFaCE และการเขียนโปรแกรม PHP เพื่อสร้างหน้าจอสำหรับรับข้อมูลเข้าจากผู้ใช้งาน และการแสดงผลพืในการสร้างแบบทดสอบตามที่ถูกผู้ใช้งานกำหนดข้อมูลตั้งต้นขึ้น รวมทั้งได้มีการประเมินค่าความถูกต้องของกลไกที่นำเสนอเพื่อยืนยันความถูกต้องของประสิทธิภาพการทำงานของกลไก โดยการเปรียบเทียบผลการประเมินระดับความยากของคำถามที่สร้างขึ้นโดยกลไกที่นำเสนอนี้กับผลการหาความยากด้วยทฤษฎีการตอบสนองข้อสอบจากการทำแบบทดสอบด้วยมนุษย์ ซึ่งพบว่ากลไกที่นำเสนอนี้สามารถให้ค่าผลลัพธ์ที่สอดคล้องตรงกับการประเมินระดับความยากด้วยมนุษย์เป็นร้อยละ 80 ดังนั้นกลไกการสร้างแบบทดสอบอัตโนมัติที่นำเสนอจึง

Received 24-08-2018

Revised 18-10-2018

Accepted 19-10-2018

สามารถนำไปประยุกต์ใช้เป็นเครื่องมือช่วยในการสร้างแบบทดสอบให้กับผู้ใช้งานได้อย่างสะดวก รวดเร็ว ลดระยะเวลาในการทำงาน รวมทั้งเพิ่มจำนวนการสร้างข้อคำถามแบบปรนัยให้มีความหลากหลาย และสามารถกำหนดระดับความยากของคำถามที่ต้องการได้ด้วย

คำสำคัญ: ระบบการสร้างแบบทดสอบอัตโนมัติ ข้อคำถามแบบปรนัย ออนโทโลยี ความคล้ายคลึงเชิงความหมาย ความสัมพันธ์เชิงความหมาย

ABSTRACT

Quiz generation system is a very important system in education. This helps to assess students' understanding of the lesson, and also allows users to easily and quickly create the test. However, the existing system of creating quizzes has limitation to the ability for creating a question, which is also entered by the user themselves. As a result, it takes time to create a test for a long time. Therefore, this research aims to provide a mechanism for the creation of automatic quiz generation mechanism for multiple choices question by applying ontological information to assess the difficulty level of the questions. The hybrid similarity was measured using a combination of semantic similarity, semantic relatedness, and property values to determine the difficulty level of the question. Then, the proposed mechanism was implemented using the RDFaCE tool and PHP program to create an interface for user input and displaying results in the creation of quiz based on user-defined data. In addition, the proposed mechanism was validated to confirm the accuracy of the mechanism's performance. By comparison, the difficulty score derived from the proposed mechanism and the item response theory. The evaluation results were consistent with 80 %. Therefore, the proposed automatic quiz generation mechanism can be applied as a tool to quickly and easily create quizzes and reduce the time of the work, including the creation of an increasing number of multiple choices question which are diverse. This method can also determine the difficulty level of the questions as required.

Keywords: Automatic quiz generation, Multiple choices question, Ontology, Semantic similarity, Semantic relatedness

1. บทนำ

ระบบการสร้างแบบทดสอบ (Quiz generation) เป็นระบบที่มีความสำคัญมากระบบหนึ่งในด้านการศึกษา เนื่องจากสามารถใช้ใน

การประเมินผลการเรียนของนักเรียนได้ว่ามีความเข้าใจในบทเรียนมากน้อยเพียงใด หรือยังขาดความรู้ความเข้าใจในด้านใดบ้าง แต่ระบบการสร้างแบบทดสอบที่มีอยู่ในปัจจุบันนั้นยังมีข้อจำกัดเรื่อง

ความสามารถในการสร้างจำนวนคำถามได้เพียงจำนวนจำกัด โดยเป็นการสร้างคำถามแบบ 1 ต่อ 1 ตามที่ผู้สอนป้อนข้อมูลเข้าสู่ระบบเท่านั้น จึงทำให้เกิดความยุ่งยาก และสิ้นเปลืองเวลาที่ใช้ในการทำงานมาก [1]

ในปัจจุบันจึงมีงานวิจัยด้านการสร้างแบบทดสอบมากขึ้น เพื่อนำไปใช้ในการเรียนการสอนและอำนวยความสะดวกให้แก่ผู้สอน แต่ระบบที่มีอยู่นั้นยังมีข้อจำกัด คือ สามารถสร้างข้อคำถามได้เพียง 2 - 3 ข้อเท่านั้น เช่น Lin C. et al. [2] นำเสนอการกำหนดระดับความยากของคำถาม โดยใช้การวัดความคล้ายคลึงเชิงความหมาย (Semantic similarity) แบบ Feature-Based ซึ่งสามารถสร้างคำถามได้เพียง 3 ข้อ Rey GÁ, et al. [3] นำเสนอวิธีการสร้างคำถามโดยใช้ความสัมพันธ์ของโหนดข้อมูล (Semantic relatedness) ซึ่งใช้การเลือกโหนดข้อมูลที่ต้องการนำมาสร้างเป็นคำถาม โดยสามารถสร้างคำถามได้เพียง 2 ข้อ จึงทำให้กลไกการสร้างคำถามที่มีอยู่ก่อนหน้านี้ไม่เพียงพอต่อการนำไปใช้งานจริงของผู้สอน

งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อนำเสนอกลไกในการสร้างคำถามอัตโนมัติประเภทคำถามแบบปรนัย (Multiple choices question: MCQ) โดยการประยุกต์ใช้ข้อมูลแบบออนโทโลยี ซึ่งกลไกที่นำเสนอได้นำวิธีการวัดค่าความคล้ายคลึงเชิงความหมาย ค่าความสัมพันธ์เชิงความหมาย และค่าคุณสมบัติของข้อมูลเข้ามาใช้งานร่วมกัน

ออนโทโลยี (Ontology) คือ การอธิบายถึงสิ่งที่เราสนใจอย่างชัดเจน โดยนำมาใช้กับเทคโนโลยีเชิงความหมาย เพื่อให้มนุษย์และคอมพิวเตอร์สามารถเข้าใจความหมายของสิ่งนั้นได้ร่วมกัน [4]

ทำให้สามารถสืบค้นค่าที่มีความหมายเหมือนกันออกมาได้แม้จะมีการใช้คำที่ต่างกัน โดยออนโทโลยีมีองค์ประกอบหลายส่วน ได้แก่ คลาส (Class), อินสแตนซ์ (Instance), ความสัมพันธ์ (Relationship), คุณสมบัติ (Property), เงื่อนไข (Constraint) เป็นต้น ซึ่งสามารถนำมาใช้ประโยชน์ในการค้นหาข้อมูลที่มีความสัมพันธ์เชิงความหมาย

การวัดค่าความคล้ายคลึงเชิงความหมาย (Semantic similarity measurement) เป็นวิธีการวัดค่าความคล้ายคลึงกันของข้อมูล โดยใช้การวัดค่าความเหมือนทางด้านโครงสร้างข้อมูลและความหมายของข้อมูล [5 - 7] ซึ่งมีวิธีการวัดค่าได้ในหลายแนวทาง เช่น แบบ Structure-based, Feature-based, Information content-based และ Hybrid

สำหรับการวัดค่าความสัมพันธ์เชิงความหมาย (Semantic relatedness) เป็นการวัดค่าความเกี่ยวข้องกันของโหนดข้อมูลที่มีการเชื่อมโยงถึงกันในรูปแบบต่าง ๆ [8] โดยโหนดข้อมูลจะมีรูปแบบของการเชื่อมโยงระหว่างกันเป็น 3 ลักษณะ ได้แก่ การเชื่อมโยงกันโดยตรง การเชื่อมโยงกันโดยอ้อม และการไม่เชื่อมโยงถึงกัน

ในส่วน ของ ค่าคุณสมบัติของข้อมูล (Property) ใช้อธิบายลักษณะของคลาสแต่ละคลาสหรืออินสแตนซ์แต่ละตัว [4]

จากข้อมูลที่กล่าวมาข้างต้น ทั้งค่าระยะห่างตำแหน่ง และคุณสมบัติของข้อมูลมีผลต่อค่าความคล้ายคลึงกันของข้อมูล ซึ่งทำให้กลไกการสร้างคำถามที่นำเสนอสามารถเพิ่มจำนวนการสร้างคำถาม และกำหนดระดับความยากง่ายของคำถามตามที่ต้องการได้

2. วิธีดำเนินการวิจัย

งานวิจัยนี้มี 2 กระบวนการ คือ การสร้างกลไกในการสร้างแบบทดสอบอัตโนมัติ และการนำกลไกที่ได้มาประเมินค่าความถูกต้อง เพื่อวัดประสิทธิภาพการทำงานของกลไกนั้น โดยมีรายละเอียด ดังนี้

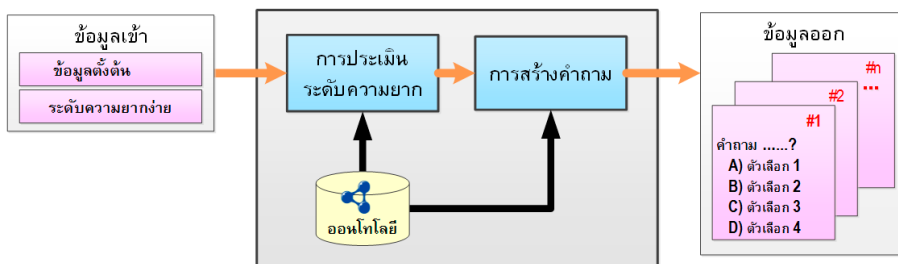
2.1 กระบวนการสร้างแบบทดสอบอัตโนมัติ

มีขั้นตอนการทำงานอยู่ 3 ขั้นตอน ได้แก่

1) การรับข้อมูลเข้า เป็นการรับข้อมูลเข้ามาจากผู้ใช้งาน 2 ข้อมูล คือ ข้อมูลที่ต้องการนำมาสร้างคำถาม และ ระดับความยากของคำถามที่ต้องการสร้าง

2) การสร้างแบบทดสอบ ประกอบด้วย 2 ขั้นตอนการทำงานย่อย คือ การประเมินระดับความยาก และการสร้างคำถาม เพื่อช่วยในการเพิ่มจำนวนการสร้างคำถามให้ได้จำนวนมากขึ้น รวมทั้งมีความหลากหลายของระดับความยากง่ายของคำถาม โดยใช้ผลการทำงานจากการประเมินระดับความยากของคำถามเข้ามาใช้งาน

3) การส่งข้อมูลออก เป็นการแสดงผลลัพธ์การสร้างคำถาม ตามข้อมูล ตั้งต้นและระดับความยากที่ผู้ใช้งานกำหนดขึ้นมาตั้งแต่แสดงแผนภาพการทำงานของกลไกในภาพที่ 1



ภาพที่ 1 กลไกการสร้างแบบทดสอบอัตโนมัติ

2.1.1 การเตรียมและรับข้อมูลเข้าจากผู้ใช้งาน

เป็นการเพิ่มข้อมูลตัวอักษรเข้าสู่ระบบด้วยผู้ใช้งานเพื่อนำมาใช้เป็นข้อมูลในการสร้างแบบทดสอบ โดยใช้โปรแกรม RDFaCE (RDFa content editor) เป็นปลั๊กอินสำหรับ TinyMCE text editor ซึ่งเป็นเครื่องมือที่ใช้ในการแก้ไขตัวอักษรหรือข้อความต่าง ๆ บนเว็บไซต์ในรูปแบบคล้ายกับการใช้งานโปรแกรม Microsoft Word ที่สามารถทำตัวหนา ตัวเอียง ชิดซ้าย ชิดขวา

ตรงกลาง ใส่ภาพ หรือใส่ตารางได้ ซึ่งจะช่วยให้สามารถทำงานกับข้อความได้สะดวกและสวยงามมากยิ่งขึ้น [9] โดยโปรแกรม RDFaCE จะเป็นการประยุกต์สำหรับแนวความคิดแบบ WYSIWYM (What you see is what you mean) ซึ่งจะช่วยให้ผู้ใช้งานสามารถมองเห็นภาพของหน้าเว็บเพจที่สร้างได้ง่ายขึ้น เพื่ออำนวยความสะดวกในการอธิบายความหมายของข้อมูลบนเว็บ และแก้ไขข้อมูล ซึ่งจะนำข้อมูลที่ได้นี้ไปเพิ่มเข้าสู่ฐานข้อมูลของโปรแกรมในการคำนวณค่าระดับความยากของตัวเลือกต่อไป

2.1.2 การสร้างแบบทดสอบ

2.1.2.1 การประเมินระดับความยากง่าย

เป็นการคำนวณหาค่าระดับความยากของตัวลอกจากที่นำมาใช้สร้างตัวเลือกในคำถามแบบปรนัย โดยการนำมาเปรียบเทียบกับคำตอบที่ถูกต้อง โดยใช้เทคนิควิธีการคำนวณหาความคล้ายคลึงแบบลูกผสม [10] ดังแสดงในสมการ (1)

$$DIF = (SIM_{HYB} \times 0.4) + (FIL \times 0.6) \quad (1)$$

โดย

DIF คือ ค่าคะแนนระดับความยากของตัวลอกจาก
 SIM_{HYB} คือ ค่าความคล้ายคลึงเชิงความหมายแบบลูกผสม โดยการเปรียบเทียบระหว่างตัวลอกจากกับคำตอบที่ถูกต้อง

FIL คือ ค่าผลการกรองข้อมูลโดยพิจารณาจากค่าคุณสมบัติของคำตอบที่ถูกต้อง

โดยค่าความคล้ายคลึงกันของข้อมูลมีค่าในช่วงตั้งแต่ 0.0 - 1.0 จึงได้มีการนำไปใช้เพื่อแบ่งช่วงของระดับความยากของคำถามออกเป็น 5 ระดับ ซึ่งหากมีค่าระดับความคล้ายคลึงกันสูง ก็จะถูกจัดให้เป็นตัวเลือกที่มีระดับความยากสูง ดังแสดงรายละเอียดในตารางที่ 1

ตารางที่ 1 ระดับความยากของคำถาม 5 ระดับ

ระดับความยาก	คะแนนความยาก (DIF)
ยากที่สุด	0.81 - 1.00
ยาก	0.61 - 0.80
ปานกลาง	0.41 - 0.60
ง่าย	0.21 - 0.40
ง่ายที่สุด	0.00 - 0.20

2.1.2.2 การสร้างคำถาม

เป็นการนำข้อมูลตั้งต้น และระดับความยากของตัวเลือกนำมาใช้สร้างเป็นคำถามขึ้นตามระดับความยากที่ต้องการ [11] โดยมีแนวทางในการสร้างคำถามใน 2 รูปแบบ คือ การเปลี่ยนคำถามและการเปลี่ยนตัวเลือกคำตอบ

ก) การเปลี่ยนคำถาม

เป็นการสร้างคำถามใหม่ขึ้นมาให้มีความแตกต่างกันโดยใช้ความสัมพันธ์ของข้อมูลตั้งต้นในการนำมาสร้างคำถามใหม่ ซึ่งมีอยู่ 3 แนวทาง ดังนี้

การสลับโหนดคำถาม เป็นการใช้อยู่ข้อมูลตั้งต้นตัวเดิม แต่มีการสลับโหนดคำถามเพื่อนำมาสร้างคำถามใหม่ เช่น “เมืองอะไรเป็นเมืองหลวงของประเทศไทย?” สลับโหนดเป็นคำถามว่า “ประเทศไทยมีเมืองหลวงคือเมืองอะไร?”

การใช้ข้อมูลที่คล้ายคลึงกับข้อมูลตั้งต้น เป็นการใช้ข้อมูลที่มีความคล้ายคลึงกับข้อมูลตั้งต้น โดยใช้คำสั่งภาษา SPARQL ในการค้นหาโหนดข้อมูลที่มีความคล้ายคลึงกัน โดยใช้ค่าคุณสมบัติของข้อมูลตั้งต้นนำมาค้นหาข้อมูลใหม่ [12] เช่น ค้นหาทุกข้อมูลที่มีค่าคุณสมบัติ :isCapitalOf จะได้ความสัมพันธ์ของข้อมูลใหม่ที่สามารถนำมาสร้างเป็นคำถามใหม่ได้

การใช้ข้อมูลที่มีความสัมพันธ์กับข้อมูลตั้งต้น เป็นการค้นหาข้อมูลที่มีความสัมพันธ์กับข้อมูลตั้งต้น โดยใช้คำสั่งภาษา SPARQL ในการค้นหาข้อมูล เช่น ค้นหาทุกข้อมูลที่มีความสัมพันธ์กับ Thailand เพื่อนำมาสร้างเป็นคำถามใหม่

ข) การเปลี่ยนตัวเลือก

เป็นวิธีการสร้างคำถามให้มีความหลากหลายมากขึ้น โดยใช้คำถามเดิม แต่เปลี่ยนแปลงตัวเลือกที่ใช้ในแต่ละข้อคำถามตาม

ระดับความยากง่ายของคำถามที่ได้มาจากผล
การประเมินระดับความยากดังที่กล่าวในหัวข้อ 2.1.2.1

2.1.3 การส่งข้อมูลออก

เป็นขั้นตอนที่กลไกนำผลที่ได้จาก
การทำงานในขั้นตอนข้างต้นนำมาแสดงผล
การสร้างคำถามตามข้อมูลและระดับความยากที่
ผู้ใช้งานกำหนดขึ้นมา โดยการแสดงคำถามที่ระบบ
สร้างขึ้นมาได้ โดยใช้โปรแกรมภาษา PHP [13] เพื่อ
แสดงผลผ่านหน้าต่าง Web user interface
เพื่อติดต่อกับผู้ใช้งาน

2.2 การประเมินค่าความถูกต้อง

การประเมินค่าความถูกต้อง (Accuracy
evaluation) เป็นวิธีการคำนวณหาค่าความถูกต้อง
ของกลไกที่นำเสนอ เพื่อใช้วัดประสิทธิภาพของกลไก
[14] ซึ่งค่าความถูกต้องของกลไกที่คำนวณได้นั้น
จะเป็นสัดส่วนของจำนวนคำถามที่มีผลการประเมิน
ระดับความยากตรง หรือใกล้เคียงกับผลการประเมิน
โดยการหาค่าความยากด้วยทฤษฎีการตอบสนอง

ข้อสอบ (Item response theory: IRT) ซึ่งเป็น
การทำแบบทดสอบจริงด้วยผู้เรียน [15] ดังแสดงสูตร
การคำนวณในสมการ (2)

$$Accuracy = \frac{R}{A} \quad (2)$$

โดยที่

Accuracy คือ ค่าความถูกต้องของกลไกที่นำเสนอ

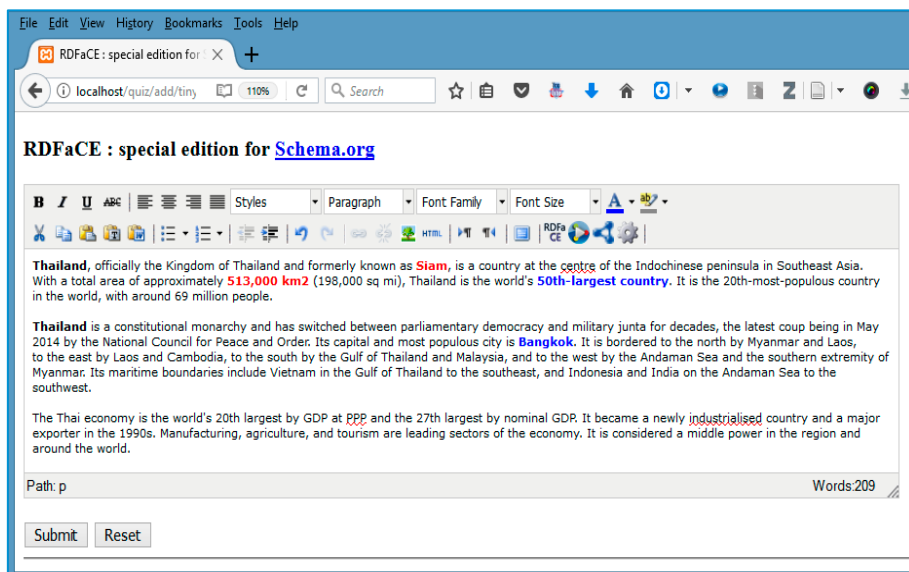
R คือ จำนวนคำถามที่มีผลการประเมินระดับ
ความยากตรงกับผลการประเมินด้วยทฤษฎี
การตอบสนองข้อสอบ

A คือ จำนวนคำถามทั้งหมด

3. ผลการวิจัยและอภิปรายผล

3.1 การนำข้อมูลเข้าจากผู้ใช้

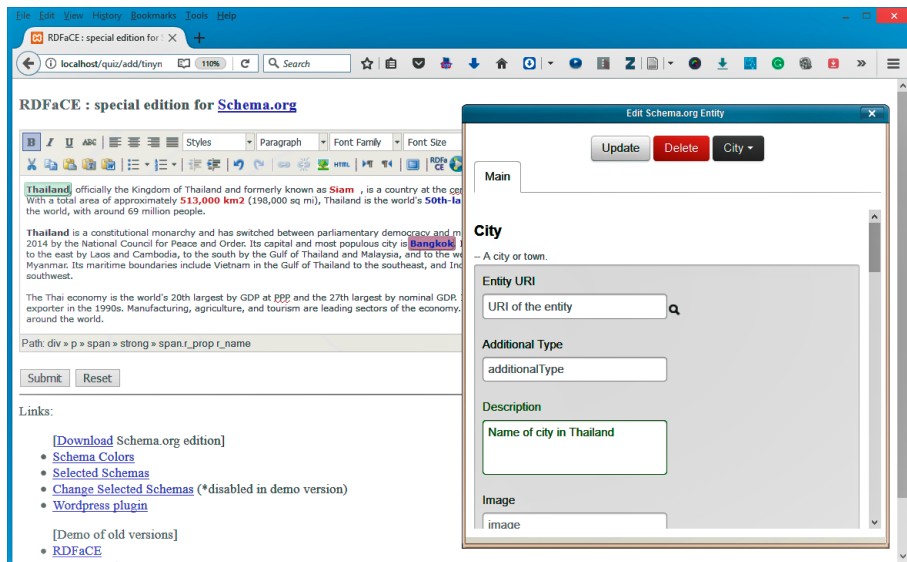
ผู้ใช้งานสามารถเพิ่มข้อมูลแบบตัวอักษร
ในเรื่องที่ต้องการสร้างแบบทดสอบเข้าสู่ระบบ
ดังแสดงในภาพที่ 2



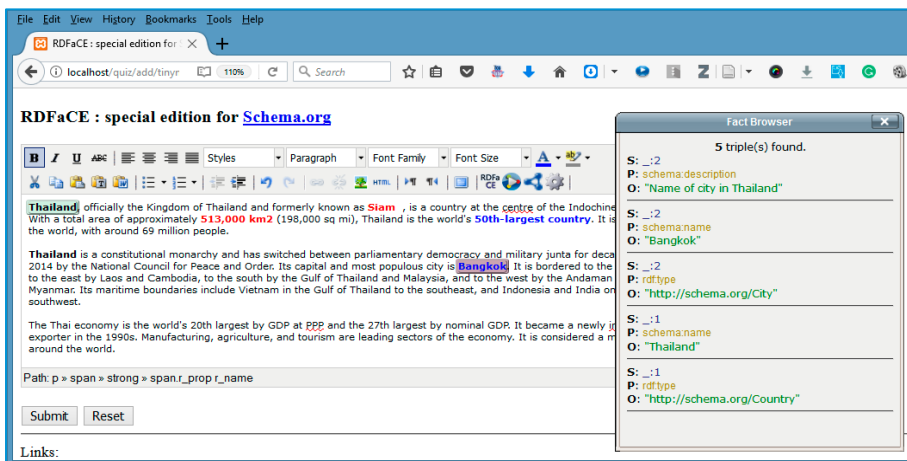
ภาพที่ 2 หน้าจอตัวอย่างการเพิ่มข้อมูลเข้าสู่ระบบ

จากนั้นผู้ใช้งานสามารถเพิ่มรายละเอียดของข้อมูลเพิ่มเติมเข้าสู่ระบบได้ โดยคลิกที่ชื่อของสิ่งนั้น และเพิ่มรายละเอียดต่าง ๆ เข้าไป ดังแสดงในภาพที่ 3

โดยระบบจะแสดงความสัมพันธ์ของข้อมูลที่เพิ่มเข้าสู่ระบบในรูปแบบ RDF triple (Subject-predicate-object) [4] ดังแสดงในภาพที่ 4

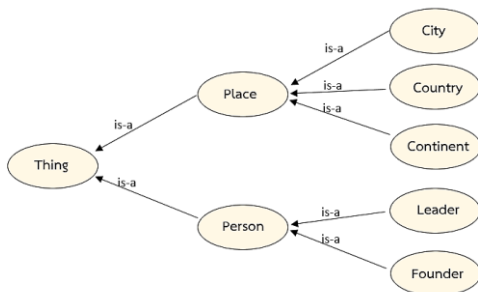


ภาพที่ 3 การเพิ่มรายละเอียดข้อมูลของสิ่งต่าง ๆ



ภาพที่ 4 หน้าจอแสดงความสัมพันธ์ของข้อมูลในรูปแบบ RDF triple

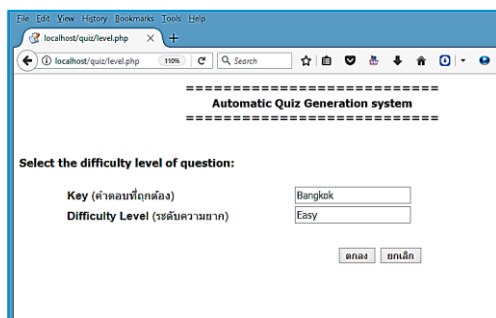
โดยข้อมูลที่ใช้เป็นตัวอย่างทดสอบนี้เป็นข้อมูลด้านภูมิศาสตร์ (Geography) ซึ่งนำไปสร้างแบบทดสอบสำหรับรายวิชาสังคมของนักเรียนชั้นประถมศึกษาปีที่ 6 โดยเป็นการเก็บข้อมูลทางภูมิศาสตร์ สังคมศาสตร์ เช่น ชื่อประเทศ ชื่อเมือง ชื่อทวีป ชื่อผู้นำ เป็นต้น ดังแสดงโครงสร้างของออนโทโลยีในภาพที่ 5



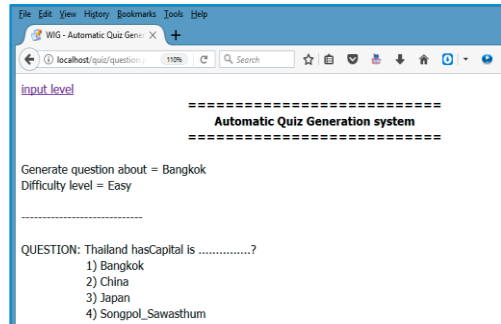
ภาพที่ 5 โครงสร้างของออนโทโลยี Geography

3.2 การสร้างคำถาม

จากการนำข้อมูลที่มีอยู่ในชุดข้อมูลมาสร้างเป็นโจทย์คำถามแบบปรนัยตามข้อมูลตั้งต้น และระดับความยากง่ายที่ผู้ใช้งานต้องการ ได้ผลลัพธ์ดังแสดงในภาพที่ 6 และ 7



ภาพที่ 6 หน้าจอแสดงการกำหนดข้อมูลตั้งต้นและระดับความยากที่ต้องการสร้างคำถาม



ภาพที่ 7 หน้าจอแสดงผลการสร้างคำถามตามข้อมูลตั้งต้น และระดับความยากของคำถามที่ต้องการ

3.3 การประเมินค่าความถูกต้อง

ในการทดสอบความถูกต้องของประสิทธิภาพการทำงานของกลไกที่นำเสนอได้ทำการเปรียบเทียบกับ การหาค่าความยากด้วยทฤษฎีการตอบสนองข้อสอบที่ได้กล่าวถึงในหัวข้อ 2.2 โดยนำผลลัพธ์ที่ได้จากการทำแบบทดสอบจริงจำนวน 10 ข้อคำถามของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 50 คน จาก 4 โรงเรียนในจังหวัดสงขลา และนำกลไกที่นำเสนอขึ้นมาใช้ประเมินระดับความยากของชุดแบบทดสอบ เพื่อเปรียบเทียบประสิทธิภาพความถูกต้องของกลไกการทำงาน ซึ่งจากผลการทดสอบได้ผลลัพธ์ที่มีค่าระดับความยากที่เหมือน หรือใกล้เคียงกันกับผลการประเมินระดับความยากด้วยกลไกที่นำเสนอจำนวน 8 ข้อ จากทั้งหมด 10 ข้อคำถาม ได้ผลการประเมินค่าความถูกต้อง เท่ากับ 0.8 หรือมีค่าความถูกต้องคิดเป็นร้อยละ 80 ของผลการประเมินระดับความยากจากคำถามทั้งหมด และเมื่อนำมาประยุกต์ใช้งานโดยการเขียนโปรแกรมเพื่อสร้างคำถามตามการกำหนดข้อมูลตั้งต้น และระดับความยากที่ผู้ใช้งานต้องการ พบว่า กลไกที่นำเสนอสามารถเพิ่มจำนวนการสร้างคำถามให้มีความหลากหลายและมากขึ้นกว่างานวิจัยเดิมที่มีอยู่

[2-3] รวมทั้งสามารถกำหนดระดับความยากง่ายของคำถามตามที่ใช้งานต้องการได้สูงถึง 5 ระดับความยากดังแสดงการเปรียบเทียบงานวิจัยที่เกี่ยวข้องในตารางที่ 2

ตารางที่ 2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง

งานวิจัย	จำนวนคำถามที่สร้างได้	วิธีการที่ใช้		
		ค่าความคล้ายคลึง	ค่าความสัมพันธ์	ค่าคุณสมบัติของข้อมูล
Lin C et al. [2]	3	✓	X	X
Rey GÁ et al. [3]	2	X	✓	X
งานที่นำเสนอ	5 ขึ้นไป	✓	✓	✓

4. สรุปผลการวิจัย

งานวิจัยนี้เป็นการนำเสนอกลไกการสร้างแบบทดสอบอัตโนมัติสำหรับคำถามแบบปรนัยด้วยวิธีการวัดค่าความคล้ายคลึงแบบลูกผสม เพื่อช่วยแก้ไขข้อด้อยในงานวิจัยด้านการสร้างแบบทดสอบเดิม ที่มีอยู่ที่สามารถสร้างแบบทดสอบได้จำนวนจำกัดเท่านั้น โดยกลไกที่นำเสนอนี้เป็นการใช้วิธีการแบบเชิงความหมายทั้งในรูปแบบความคล้ายคลึงเชิงความหมาย ความสัมพันธ์เชิงความหมาย ร่วมกับการใช้ค่าคุณสมบัติของข้อมูล โดยงานวิจัยนี้พบว่า

1) การนำเทคนิควิธีการวัดค่าความคล้ายคลึงเชิงความหมาย และค่าความสัมพันธ์เชิง

ความหมายเข้ามาใช้งานร่วมกันแบบลูกผสมสามารถนำมาใช้สร้างแบบทดสอบอัตโนมัติที่มีระดับความยากได้ละเอียดมากขึ้นถึง 5 ระดับความยาก

2) การปรับปรุงกลไกการสร้างคำถามโดยใช้ประโยชน์จากแนวทางการประเมินระดับความยากแบบลูกผสมร่วมกับการสลับโหนดข้อมูลเพื่อเปลี่ยนบทบาทของประธาน และกรรมในโจทย์คำถามสามารถเพิ่มจำนวนการสร้างข้อคำถามแบบปรนัยให้ได้จำนวนมากขึ้นได้

3) กลไกที่นำเสนอนี้เมื่อนำไปประเมินค่าความถูกต้องด้วยการเปรียบเทียบกับผลการหาค่าความยากด้วยทฤษฎีการตอบสนองข้อสอบด้วยผู้เรียนจริง เพื่อยืนยันประสิทธิภาพการทำงานของกลไก พบว่ามีค่าความถูกต้องตรงกันถึงร้อยละ 80

สำหรับข้อจำกัดของกลไกการสร้างคำถามแบบอัตโนมัติที่พัฒนาขึ้น ซึ่งควรศึกษา และพัฒนางานวิจัยต่อยอดเพิ่มเติม ดังต่อไปนี้

1) ควรปรับปรุงความสามารถของกลไกประเมินระดับความยากแบบลูกผสมให้รองรับชุดข้อมูลที่มีขนาดใหญ่ และไม่ทราบจำนวนโหนดข้อมูลทั้งหมดที่เกี่ยวข้องได้ เช่น ข้อมูลเว็บแบบเชิงความหมายด้วยกลไกทำงานแบบ Linked data ซึ่งมีความแตกต่างไปจากสิ่งแวดล้อมในการทำงานของกลไกที่นำเสนอ เนื่องจากข้อมูลที่ใช้ในกลไกที่นำเสนอต้องเป็นชุดข้อมูลที่มีขอบเขตชัดเจนและทราบจำนวนโหนดที่แน่นอน

2) ควรปรับปรุงรูปประโยคของโจทย์คำถามแบบปรนัยที่สร้างขึ้นโดยอัตโนมัติให้มีความสละสลวยมากขึ้น ซึ่งสามารถทำได้โดยใช้วิธีการประมวลผลภาษาธรรมชาติเข้ามาพิจารณาร่วมด้วย ซึ่งยังไม่ได้ดำเนินการในงานวิจัยนี้

5. อ้างอิง

- [1] Fattoh IE, Aboutabl AE, Haggag MH. Semantic Attributes Model for Automatic Generation of Multiple Choice Questions. *Int J Comput Appl* 2014; 103(1):18–24.
- [2] Lin C, Liu D, Pang W, et al. Automatically Predicting Quiz Difficulty Level Using Similarity Measures. *K-CAP 2015. The 8th International Conference on Knowledge Capture*; 2015 Oct 7-10; Palisades, NY, USA. New York: ACM; 2015. P. 1:1–1:8.
- [3] Rey GÁ, Celino I, Alexopoulos P, et al. Semi-Automatic Generation of Quizzes and Learning Artifacts from Linked Data. *The 2nd International Workshop on Learning and Education with the Web of Data at the World Wide Web Conference 2012 (WWW2012)*; 2012 Apr 17; Lyon, France. New York: ACM; 2012. P. 1-3.
- [4] งามนิจ อัจฉรินทร์. เทคโนโลยีเว็บแบบสื่อความหมาย (Semantic Web Technologies). พิมพ์ครั้งที่ 3. ขอนแก่น: คลังน่านวิทยา; 2557.
- [5] Elavarasi S, Akilandeswari J, Menaga K. A Survey on Semantic Similarity Measure. *Int J Res in Advent Technol*. 2014; 2:389–98.
- [6] Abdelrahman AMB, Kayed A. A Survey on Semantic Similarity Measures between Concepts in Health Domain. *Am J Comput Math*. 2015; 5(2):204-14.
- [7] Gan M, Dou X, Jiang R. From Ontology to Semantic Similarity: Calculation of Ontology-Based Semantic Similarity. *Sci World J*. 2013; 2013:1-11.
- [8] Meng L, Huang R, Gu J. A Review of Semantic Similarity Measures in WordNet. *Int J Hybrid Inf Technol*. 2013; 6(1):1–12.
- [9] Agile Knowledge Engineering and Semantic Web (AKSW). [Internet]. Germany: RDFaCE; 2018 [cited 2018 Aug 26]; [about 3 screens]. Available from: <http://aksw.org/Projects/RDFaCE.html>
- [10] Nuthong S, Witosurapot S. Enabling Fine Granularity of Difficulty Ranking Measure for Automatic Quiz Generation. *The 9th International Conference on Information Technology and Electrical Engineering (ICITEE 2017)*; 2017 Oct 12-13; Centara Grand Beach Resort Phuket. Bangkok: The University; 2017. P. 227-32.
- [11] ศศิธร หนูทอง, สุนทร วิฑูรพจน์. การปรับปรุงกลไกทำงานแบบชาญฉลาดสำหรับการสร้างแบบคำถามโดยอัตโนมัติ. *การประชุมทางวิชาการระดับชาติด้านเทคโนโลยีคอมพิวเตอร์และระบบสารสนเทศประยุกต์ ครั้งที่ 12*; 19-21 ก.ค. 2560; มหาวิทยาลัยเทคโนโลยีราชมงคลกรุงเทพ. กรุงเทพฯ: มหาวิทยาลัย; 2560. น. 190-4.

- [12] World Wide Web Consortium (W3C). [Internet]. Beihang: SPARQL Query Language for RDF; 2013 [cited 2018 Aug 26] ; [about 87 screens]. Available from: <https://www.w3.org/TR/rdf-sparql-query>
- [13] php.net [Internet]: The PHP Group; 2018 [cited 2018 Aug 26]. Available from: <http://www.php.net>
- [14] Kaur J, Gupta V. Effective Question Answering Techniques and their Evaluation Metrics. Int J Comput Appl. 2013; 65(12):30-7.
- [15] สัจจวรรณ จัดกระโทก. ทฤษฎีการตอบข้อสอบ Item Response Theory [อินเทอร์เน็ต]. กรุงเทพฯ: สำนักงานคณะกรรมการวิจัยแห่งชาติ; 2556 [เข้าถึงเมื่อ 26 ส.ค. 2561]. จ าก : http://www.priv.nrct.go.th/ewt_dl.php?nid=1057