การพัฒนาโมเดลสำหรับจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา

นายนพพล ตั้งสุภาชัย

มหาวิทยาลัยเทคโนโลยีสุรนารี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาการสารสนเทศดุษฎีบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2558

THE DEVELOPMENT OF LEARNING OBJECTS STORAGE AND RETRIEVAL MODEL FROM OPEN DATA

Noppol Thangsupachai

มหาวิทยาลัยเทคโนโลยีสุรนารี

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Information Science in Information Technology Suranaree University of Technology

Academic Year 2015

การพัฒนาโมเดลสำหรับจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา

มหาวิทยาลัยเทคโนโลยีสุรนารีอนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์

_____

(ศาสตราจารย์ ดร.วิลาศ วูวงศ์)
ประธานกรรมการ

_____

(ผู้ช่วยศาสตราจารย์ ดร.ศุภกฤษฎิ์ นิวัฒนากูล)
กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)

_____

(รองศาสตราจารย์ เภสัชกร ดร.อนุชัย ธีระเรืองไชยศรี)
กรรมการ

_____

(รองศาสตราจารย์ ดร.นิตยา เกิดประสพ)
กรรมการ

_____

(ผู้ช่วยศาสตราจารย์ ดร.ธรา อั่งสกุล)
กรรมการ

_____

(อาจารย์ ดร.นิศาชล จำนงศรี)
กรรมการ

_____

(ศาสตราจารย์ ดร.ชูกิจ ลิมปิจำนงค์)
รองอธิการบดีฝ่ายวิชาการและนวัตกรรม

_____

(รองศาสตราจารย์ ดร.วีรพงษ์ พลนิกรกิจ)
คณบดีสำนักวิชาเทคโนโลยีสังคม

มหาวิทยาลัยเทคโนโลยีสุรนารี

นพพล ตั้งสุภาชัย : การพัฒนาโมเดลสำหรับจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา (THE DEVELOPMENT OF LEARNING OBJECTS STORAGE AND RETRIEVAL MODEL FROM OPEN DATA) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร. ศุภกฤษฎิ์ นิวัฒนากูล, 224 หน้า.

งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา โดยกระบวนการวิจัยออกเป็น 3 ส่วน ประกอบด้วย (1) การศึกษาถึงพฤติกรรมการสืบค้นและเลือกใช้เลิร์นนิงอ็อบเจ็กต์จากอินเทอร์เน็ต (2) การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ และ (3) การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

ส่วนที่ 1 เก็บข้อมูลจากแบบสอบถามที่รวบรวมจากครูผู้สอนในหลักสูตรสองภาษา (English Program) ของโรงเรียนในสังกัดของ สพฐ. และโรงเรียนเอกชน รวม 195 ชุด พบว่าเครื่องมือที่นิยมใช้ในการสืบค้น เลิร์นนิงอ็อบเจกต์คือ เสิร์ชเอนจิน ร้อยละ 34 โดยนิยมค้นจากคำสำคัญและชื่อเรื่องมากที่สุด ร้อยละ 26.7 เนื้อหาที่ต้องการนำไปใช้ในการเรียนการสอนมากที่สุดคือ เนื้อหาบรรยาย ร้อยละ 25.4 และแบบฝึกหัด ร้อยละ 23.8 โดยนิยมนำไปใช้เป็นส่วนเสริมในการเรียนการสอนมากกว่าเป็นเนื้อหาหลัก ร้อยละ 27.5 ปัญหาของการสืบค้นได้แก่ การใช้ภาษาอังกฤษในการสืบค้น แหล่งข้อมูลมีลิขสิทธิ์หรือมีค่าใช้จ่ายในการเข้าถึง ผลการสืบค้นมีความซ้ำซ้อนหรือไม่มีอยู่จริง เนื้อหาไม่ตรงตามคำสำคัญ และแหล่งข้อมูลไม่น่าเชื่อถือ ข้อมูลที่ใช้เพื่อพิจารณาคัดเลือกผลการสืบค้นสองอันดับแรกได้แก่ ชื่อเรื่อง และประเภทของไฟล์ รองลงไปคือ รายละเอียดเนื้อหา หัวเรื่อง ภาษา แหล่งที่มาของข้อมูล ระดับความยากง่ายของเนื้อหา และหน่วยงานที่เผยแพร่ข้อมูล

ส่วนที่ 2 การออกแบบโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ พบว่า โมเดลการจัดเก็บเลิร์นนิง อ็อบเจกต์ซึ่งประกอบด้วยโครงสร้างข้อมูล 10 คลาส เพื่อรองรับข้อมูลเมทาดาทาที่ใช้มาตรฐานต่างกันได้ โดยใช้เทคนิควิธีการผสานเค้าร่างเมทาดาทาด้วยการสร้างคลังคำศัพท์ของเมทาดาทาจากเวิร์ดเน็ต และเทคนิคการวัดความคล้ายคลึงของเมทาดาทาแต่ละองค์ประกอบ ใช้เครื่องมือ D2RQ ช่วยในการผสานเมทาดาทา และสุดท้ายทำการแปลงเอกสารให้อยู่ในโครงสร้างแบบเปิดเสรีเพื่อรองรับการแลกเปลี่ยนข้อมูลแบบเสรีได้ สำหรับส่วนโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์จะทำการเตรียมข้อมูลก่อนทำการสืบค้น โดยแปลงข้อมูลอาร์ดีเอฟให้มีรูปแบบเป็นตารางด้วยเทคนิคแบบคีย์-แวลู (Key-Value) เพื่อลดเวลาการเข้าถึงข้อมูล การสืบค้นเลิร์นนิงอ็อบเจกต์จะใช้ชุดคำศัพท์ที่สร้างจากคลังคำทางคณิตศาสตร์จำนวน 300 คำ โดยกำหนดโครงสร้างความสัมพันธ์ของ

คำศัพท์ตามโครงสร้างของเวิร์ดเน็ต และหนังสือเรียนรายวิชาคณิตศาสตร์ตามหลักสูตรแกนกลางของ กระทรวงศึกษาธิการ และทำการวัดความคล้ายคลึงเชิงความหมาย (Semantic Similarity Score) ประกอบด้วยการวัดระยะทางเชิงความหมาย โดยใช้ข้อมูลจากพร็อพเพอร์ตีในการกำหนดกลุ่มคำศัพท์ที่เกี่ยวข้อง เพื่อกำหนดขอบเขตในการสืบค้น และการหาคะแนนจากการวัดความคล้ายคลึงเชิงความหมายจากคำสำคัญและข้อมูลในคลาสที่มีค่าความคล้ายคลึงใกล้เคียงกัน โดยจะนำคะแนนทั้งสองส่วนมาประมวลผลเพื่อให้ค่าน้ำหนัก และในขั้นสุดท้ายจะทำการให้น้ำหนักขององค์ประกอบย่อยเมทาดาทาเพื่อจัดลำดับการผลการค้น (Ranking) ตามความต้องการของผู้ใช้

ส่วนที่ 3 ผลการประเมินประสิทธิภาพของโมเดลการจัดเก็บและสืบค้นเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทาที่พัฒนาขึ้น สามารถวัดประสิทธิภาพความถูกต้องในการค้นคืนเลิร์นนิงอ็อบเจกต์จากค่าความแม่นยำ (Precision) ร้อยละ 95.55 ค่าความระลึก (Recall) ร้อยละ 88.91 และค่าอัตราการรู้จำ (F-Measure) ร้อยละ 91.92 รวมถึงได้ทดสอบความเร็วในการสืบค้นข้อมูลโดยเปรียบเทียบกับเจนาเฟรมเวิร์ค (Jena Framework) พบว่าระบบที่พัฒนาขึ้นสามารถสืบค้นข้อมูลที่มีขนาดแตกต่างกันจะใช้เวลาการสืบค้นใกล้เคียงกัน โดยข้อมูลขนาดใหญ่ขึ้นจะใช้เวลาเพิ่มขึ้นเพียงเล็กน้อย แต่คำสั่งเจนาจะใช้เวลาในการสืบค้นเพิ่มขึ้นแบบแปลผันตามขนาดของชุดข้อมูล ทั้งนี้เมื่อนำไปประเมินประสิทธิภาพในการใช้ในการสืบค้นเลิร์นนิงอ็อบเจกต์โดยผู้ใช้ ปรากฏว่าได้รับการประเมินโดยรวมในระดับดีมาก

มหาวิทยาลัยเทคโนโลยีสุรนารี

สาขาวิชาเทคโนโลยีสารสนเทศ ลายมือชื่อนักศึกษา_____

ปีการศึกษา 2558 ลายมือชื่ออาจารย์ที่ปรึกษา_____

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม_____

NOPPOL THANGSUPACHAI : THE DEVELOPMENT OF LEARNING OBJECTS STORAGE AND RETRIEVAL MODEL FROM OPEN DATA. THESIS ADVISOR : ASST. PROF. SUPHAKIT NIWATTHANAKUL, Ph.D. 224 PP.

SEARCHING BEHAVIOUR/ LEARNING OBJECT/ METDATA MAPPING/ SEMANTIC SEMILARITY SCORE/ RANKING

This research aimed to design and develop model of learning object storage and retrieval from open data. The process of research was divided into 3 parts which were 1) Studying behavior of searching and selecting learning object from internet 2) Designing and developing model of learning object storage and retrieval 3) Evaluating model of learning object storage and retrieval.

1. The research was collected data from questionnaires which were gathered from teachers in English program of schools which belong to Office of the Basic Education Commission and private schools. The numbers of questionnaires were 195 sets; the result was found that popular tool which was used to search learning object was search engine 34%. They most liked to search from keyword and subject 23.8%. The contents that were most required to use for schooling were description 25.4% and example 23.8%. These contents were used as additional lessons more than main content 27.5%. The problem of searching was using English for searching, resource had copyright or additional expenses for access, the result of searching was duplicated or unreal, content did not serve the keyword and resources were not reliable. The information which was used to consider and select result of searching were subject and type of file and details of content, topic, language, resource, level of complication of content and department that published information.

2. Designing and developing model of learning object storage and retrieval were found that model of learning object storage which consisted of information structure 10 classes to serve metadata with different standards. The technique of building metadata by creating vocabulary's groups of metadata from technique of measurement of similarity of each element of metadata were used. Finally, the document was changed to be opened structure to serve exchanging information freely. The model of learning object storage and retrieval will prepare information before searching; RDF data will be transformed to table by Key-Value's technique. To search learning object used sets of vocabulary and Semantic Similarity Score have measured by using information from data property to specify scope of searching. The scores were brought to evaluate to give results weighted in ranking list.

3. Result of evaluating model of learning object storage and retrieval from open data which was developed. It could measure the efficiency of correctness from learning object retrieval from precision 95.55%, recall 88.91% and F-measure 91.92% and the speed of searching was tested by comparing with Jena Framework. It was found that the developed system was faster than Jana when search information which had different sizes and took similar time. The larger information took a little bit more time. However, the Jena took more time for searching as per the size of dataset. When it was evaluated efficiency for searching learning object by users, it was found that overall of evaluation was very good.

School of Information Technology Student's Signature_____

Academic Year 2015 Advisor's Signature_____

Co-Advisor's Signature_____

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบคุณบุคคลและกลุ่มบุคคลต่อไปนี้ ที่ได้กรุณาให้ความช่วยเหลือ ให้คำปรึกษา ให้คำแนะนำ และเป็นกำลังใจให้ตลอดระยะเวลาในการเรียน และการทำวิจัยในครั้งนี้ โดยความสำเร็จในครั้งนี้จะเกิดขึ้นไม่ได้เลย ถ้าไม่ได้รับการสนับสนุนจากบุคคลดังต่อไปนี้

ขอขอบคุณ มหาวิทยาลัยเทคโนโลยีสุรนารีที่ให้การสนับสนุนทุนการศึกษา ทุนสนับสนุนการทำวิจัย และการนำเสนอผลงานวิชาการจนสำเร็จการศึกษาตามหลักสูตรนี้

ขอขอบพระคุณอาจารย์ที่ปรึกษาทั้งสองท่าน ผู้ช่วยศาสตราจารย์ ดร.ศุภกฤษฎิ์ นิวัฒนากูล และอาจารย์ ดร.นิศาชล จำนงศรี ที่คอยแนะนำ เป็นกำลังใจ ให้ข้อคิด และช่วยผลักดันให้งานวิจัยนี้สำเร็จไปด้วยดี

ขอขอบคุณคณะกรรมการสอบวิทยานิพนธ์ ศาสตราจารย์ ดร.วิลาศ วูวงศ์ ผู้ทำหน้าที่ประธานกรรมการ ผู้ช่วยศาสตราจารย์ ดร.อนุชัย ธีระเรืองไชยศรี รองศาสตราจารย์ ดร.นิตยา เกิดประสพ และผู้ช่วยศาสตราจารย์ ดร.ธรา อั่งสกุล ผู้ทำหน้าที่กรรมการสอบวิทยานิพนธ์ทุกท่าน ที่กรุณาให้คำแนะนำ ให้คำปรึกษา และตรวจทานเนื้อหาวิทยานิพนธ์จนเสร็จสมบูรณ์

ขอขอบคุณบุคลากรประจำสาขาวิชาเทคโนโลยีสารสนเทศและบุคลากรประจาสำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารีที่ให้ความช่วยเหลือ สนับสนุนและอำนวยความสะดวกตลอดระยะเวลาการศึกษาวิจัยนี้

ขอขอบคุณ ผู้เชี่ยวชาญ รองศาสตราจารย์ ดร.สงวน วงษ์ชวลิตกุล จากมหาวิทยาลัยวงษ์ชวลิตกุล ดร.บัณฑิต วรรณประพันธ์ นักวิชาการสาธารณสุขชำนาญการพิเศษ คุณนันทนา แต้ประเสริฐ นักวิชาการสาธารณสุขชำนาญการพิเศษ และคณะ ที่ช่วยตรวจสอบ ประเมินแบบสอบถาม ให้คำแนะนำในกระบวนการวิจัยเชิงปริมาณ การเก็บแบบสอบถาม และการประมวลผลข้อมูล

ขอขอบคุณคณาจารย์ ครูผู้สอนจากโรงเรียนราชสีมาวิทยาลัย โรงเรียนสุรนารีวิทยา และจากโรงเรียนต่าง ๆ ในประเทศไทย ซึ่งกรุณาตอบแบบสอบถามและให้การสัมภาษณ์ และประเมินแบบจำลองในงานวิจัยนี้ เพื่อนำข้อมูลมาพัฒนาแบบจำลองจนเสร็จสมบูรณ์

ขอขอบคุณ ดร.ศุภชานันท์ วนภู ที่คอยช่วยเหลือ ประสานงาน และห่วงใยถามไถ่ถึงรุ่นน้องๆ เสมอ อีกทั้งเพื่อนร่วมรุ่น ผู้ช่วยศาสตราจารย์ ดร.หนึ่งหทัย ขอผลกลาง อาจารย์ ดร.พิชญสินี กิจวัฒนาถาวร ดร.คมคิด ชัชราภรณ์ นางสาวรุ่งกานต์ มูสโกภาส และนายนโรดม กิตติเดชานุภาพ

และเพื่อนสมาชิกร่วมหลักสูตรวิทยาการสารสนเทศดุษฎีบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีสุรนารี ทุกท่านที่คอยช่วยเหลือแบ่งปัน และปรับทุกข์ในยามท้อแท้ แก้ปัญหากับการเรียนที่เป็นรุ่นบุกเบิก พร้อมด้วยความรู้สึกที่ดีต่อกันเสมอมา ทำให้การศึกษาครั้งนี้สำเร็จลุล่วงไปด้วยดี

สุดท้ายนี้ ขอขอบพระคุณ นายแพทย์สมชาย ตั้งสุภาชัย ผู้เป็นบิดา นางลออศรี ตั้งสุภาชัย ผู้เป็นมารดาที่รักยิ่ง ญาติพี่น้องทุกท่าน ผู้ที่ให้การสนับสนุนทุนการศึกษา กำลังใจ ให้คำแนะนำแนวทางในการแก้ปัญหา ด้านการเรียน อีกทั้งยังให้ความช่วยเหลือในระหว่างการเรียนและทำวิจัย และบุคคลสำคัญอีกท่านคือ นางสุรัสวดี ตั้งสุภาชัย ซึ่งเป็นผู้ที่อยู่เคียงข้างเสมอมาทั้งในยามสุขและยามทุกข์ จะเป็นกำลังใจให้เสมอมา รวมถึงการดูแล ช่วยเหลือในทุก ๆ ด้าน จึงทำให้งานวิจัยและการศึกษาในครั้งนี้สำเร็จลุล่วงไปได้ด้วยดี

นพพล ตั้งสุภาชัย

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย) ..... ก
บทคัดย่อ (ภาษาอังกฤษ) ..... ค
กิตติกรรมประกาศ ..... จ
สารบัญ ..... ช
สารบัญตาราง ..... ฎ
สารบัญรูป ..... ฌ
บทที่
1 บทนำ 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์การวิจัย 7
1.3 คำถามนำการวิจัย 7
1.4 ขอบเขตการวิจัย 7
1.5 ประโยชน์ที่คาดว่าจะได้รับ 8
1.6 คำอธิบายศัพท์ 8
2 ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง 11
2.1 พฤติกรรมการสืบค้นเลิร์นนิงอ็อบเจกต์จากอินเทอร์เน็ต 12
2.1.1 ความหมายของพฤติกรรมแสวงหาสารสนเทศ 12
2.1.2 แบบจำลองวิเคราะห์กระบวนและปัจจัยของพฤติกรรมการแสวงหาสารสนเทศ 13
2.1.3 แหล่งข้อมูลในการสืบค้นเลิร์นนิงอ็อบเจกต์ 17
2.2 เลิร์นนิงอ็อบเจกต์ 23
2.2.1 ความหมายของเลิร์นนิงอ็อบเจกต์ 23
2.2.2 คุณลักษณะของเลิร์นนิงอ็อบเจกต์ 26
2.2.3 องค์ประกอบของเลิร์นนิงอ็อบเจกต์ 27

สารบัญ (ต่อ)

หน้า

2.3 การจัดเก็บเลิร์นนิงอ็อบเจกต์ 31
2.3.1 เมทาดาทา 32
2.3.2 คลังคำศัพท์สำหรับการสืบค้นสารสนเทศ 38
2.3.3 ลิงค์โอเพนดาทา (Linked Open Data) 45
2.3.4 การผสานเมทาดาทา (Metadata Mapping) 47
2.4 การค้นคืนเลิร์นนิงอ็อบเจกต์จากโอเพนดาทา 53
2.4.1 ภาษาและเครื่องมือสำหรับสืบค้นลิงค์โอเพนดาทา (Linked Open Data Searching and Tools) 54
2.4.2 การวัดความคล้ายคลึงของผลการสืบค้น (Similarity Measurement) 62
2.4.3 การจัดลำดับผลการสืบค้น (Search Result Ranking) 70
2.5 งานวิจัยที่เกี่ยวข้อง 76
2.5.1 สรุปงานวิจัยที่เกี่ยวข้อง 76
2.5.2 บทสรุปและเปรียบเทียบงานวิจัย 78
2.5.3 กรอบแนวคิดการวิจัย 79
3 วิธีดำเนินการวิจัย 80
3.1 การศึกษากระบวนการการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ 80
3.1.1 การศึกษากระบวนการการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ 80
3.1.2 การศึกษาเทคนิควิธีการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา 81
3.1.3 การศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน 81
3.2 การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา 86
3.2.1 การออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ 86
3.2.2 การออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ 198
3.3 การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ 106
3.3.1 การประเมินประสิทธิภาพของระบบโดยการทดสอบ 106

สารบัญ (ต่อ)

หน้า

3.3.2 การประเมินความสามารถในการทำงานของระบบโดยผู้เชี่ยวชาญ 108

4 ผลการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ บนอินเทอร์เน็ตของผู้ใช้ 110

4.1 ผลการวิจัย 111

4.2 สรุปและอภิปรายผล 122

4.2.1 แหล่งข้อมูลหลักที่ใช้ในการค้นหาข้อมูลเพื่อใช้ประกอบการเรียนการสอน 122

4.2.2 วิธีการสืบค้นและคัดเลือกรายการเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตเพื่อนำไปใช้ในการเรียนการสอน 123

4.2.3 การนำเสนอเลิร์นนิงอ็อบเจกต์ที่ได้เพื่อไปใช้ในการเรียนการสอน 123

4.2.4 ปัญหาที่พบในการสืบค้นเลิร์นนิงอ็อบเจกต์ 124

4.2.5 การออกแบบชุดเมทาดาทาสำหรับอธิบายเลิร์นนิงอ็อบเจกต์ 126

5 การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา 129

5.1 ผลการออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ 129

5.1.1 การออกแบบเมทาดาทาเพื่อการจัดเก็บเลิร์นนิงอ็อบเจกต์ 130

5.1.2 การผสานเมทาดาทาและการแปลงโครงสร้างอาร์ดีเอฟ 138

5.2 ผลการออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ 144

5.2.1 การออกแบบการค้นคืนเชิงความหมาย 145

5.2.2 การจัดลำดับการนำเสนอผลการสืบค้น 157

5.3 ผลการประเมินแบบจำลองการค้นคืนเลิร์นนิงอ็อบเจกต์ 161

5.3.1 ผลการประเมินประสิทธิภาพของระบบโดยการทดสอบการทำงานของและการเปรียบเทียบความเร็วในการค้นคืนเลิร์นนิงอ็อบเจกต์กับระบบอื่น 162

5.3.2 การประเมินความสามารถของระบบโดยผู้ใช้ 167

5.4 สรุปผลการออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา 175

6 สรุปผลการวิจัยและข้อเสนอแนะในการทำวิจัย 178

สารบัญ (ต่อ)

หน้า

6.1 สรุปผลการวิจัย 178
6.1.1 การศึกษากระบวนการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ 178
6.1.2 ผลการออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ 181
6.1.3 การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ 184
6.2 การประยุกต์ใช้ผลการวิจัย 185
6.3 ข้อจำกัดและข้อเสนอแนะในการทำวิจัย 186
6.3.1 ข้อจำกัดในการวิจัย 186
6.3.2 ข้อเสนอแนะในการวิจัยครั้งต่อไป 187
รายการอ้างอิง 188
ภาคผนวก
ภาคผนวก ก.แบบสอบถามการค้นคืนและเลือกใช้ข้อมูลจากแหล่งสารสนเทศบนอินเทอร์เน็ต 209
ภาคผนวก ข.แบบประเมินประสิทธิภาพแบบจำลองการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ 219
ประวัติผู้วิจัย 224

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญตาราง

ตารางที่ **หน้า**

2.1 เปรียบเทียบคุณสมบัติของแหล่งข้อมูลแบบระบบปิดและระบบเปิดเสรี 21
2.2 รูปแบบโครงสร้างประโยคคำถามสปาเคิล 54
2.3 เปรียบเทียบตัวชี้วัดเบราเซอร์สำหรับลิงค์โอเพนดาทา 56
2.4 เปรียบเทียบฟังก์ชั่นของเครื่องมือสืบค้นเชิงความหมายสำหรับลิงค์โอเพนดาทา 58
2.5 ส่วนต่อประสานภาษาธรรมชาติสำหรับการสืบค้นเชิงความหมาย 61
2.6 การจัดรูปแบบของเอกสารแบบถุงของคำ 66
2.7 บทสรุปงานวิจัยที่เกี่ยวข้อง 78
3.1 ตัวอย่างการเก็บข้อมูลแบบทริเปิลในตารางข้อมูลทั่วไป 99
3.2 การเก็บข้อมูลในแบบคีย์-แวลู คอลเล็กชั่นโหนดยอด (Vertex Collection) 100
3.3 การเก็บข้อมูลในแบบคีย์-แวลู คอลเล็กชั่นพร็อพเพอร์ตี (Predicate Collection) 100
3.4 การเก็บข้อมูลอาร์ดีเอฟ (RDFData) 100
4.1 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามประเภทของโรงเรียน 111
4.2 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามสถานะของผู้สอน 111
4.3 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามประสบการณ์การสอน 112
4.4 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามหลักสูตรที่รับผิดชอบ 112
4.5 แหล่งข้อมูลหลักที่ใช้ในการค้นหาข้อมูลเพื่อใช้ประกอบการเรียนการสอน 113
4.6 เครื่องมือที่นิยมใช้เพื่อชี้แหล่งข้อมูลในการเข้าถึงเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต 113
4.7 ข้อมูลที่นิยมใช้ในการสืบค้นข้อมูลบนอินเทอร์เน็ต 114
4.8 ลักษณะของข้อมูลเกี่ยวกับวิชาคณิตศาสตร์ที่มักจะค้นหา 115
4.9 การนำข้อมูลที่สืบค้นคืนได้ไปใช้ในการเรียนการสอนวิชาคณิตศาสตร์ 116
4.10 การตระหนักรู้ระดับของปัญหาในการสืบค้นเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต 117
4.11 ข้อมูลที่ใช้ในการพิจารณาเพื่อเปิดใช้รายการผลการสืบค้นจากหน้ารวมผลการค้น 119
4.12 การคัดกรอง หรือปรับแต่งการแสดงผลการสืบค้น 120

สารบัญตาราง (ต่อ)

ตารางที่ **หน้า**

4.13 วิธีการจัดเก็บเลิร์นนิงอ็อบเจกต์ที่มีอยู่เดิมหรือสร้างขึ้นใหม่ 121
4.14 เมทาดาทาที่ใช้ในการอธิบายเลิร์นนิงอ็อบเจกต์ 126
5.1 คลาสผู้สร้างผลงาน (Contributor) 131
5.2 คลาสแผนกและส่วนงาน (Department) 131
5.3 คลาสหน่วยบทเรียน (Lesson) 132
5.4 คลาสวัตถุประยุกต์ (ApplicationObject) 133
5.5 คลาสวัตถุสารสนเทศ (InformationObject) 133
5.6 คลาสรายการข้อมูลย่อย (Asset) 134
5.7 คลาสคำสำคัญ (Keyword) 135
5.8 คลาสรูปแบบการมีปฏิสัมพันธ์ (InteractiveType) 135
5.9 คลาสประเภทเลิร์นนิงอ็อบเจกต์ (LOType) 135
5.10 คลาสภาษา (Language) 135
5.11 ตัวอย่างข้อมูลทริเปิลของเลิร์นนิงอ็อบเจกต์ 148
5.12 คอลเล็กชันจุดยอด (Vertex Collection) 149
5.13 คอลเล็กชันคุณสมบัติของทรัพยากร (Predicate Collection) 150
5.14 คอลเล็กชันอาร์ดีเอฟเลิร์นนิงอ็อบเจกต์ (RDF Collection of Learning Object) 150
5.15 คอลเล็กชันจุดยอดที่เกี่ยวข้องกับคำค้น 152
5.16 คอลเล็กชันคุณสมบัติของทรัพยากรที่เกี่ยวข้องกับคำค้น 152
5.17 ตัวอย่างการวัดค่าความคล้ายคลึงที่ได้จากการสืบค้นเลิร์นนิงอ็อบเจกต์ 160
5.18 เปรียบเทียบผลการวัดความคล้ายคลึงแบบมีการเพิ่มค่าน้ำหนักและไม่เพิ่มค่าน้ำหนัก 160
5.19 ข้อมูลทั่วไปของผู้ตอบแบบประเมินความสามารถระบบสืบค้น 172
5.20 ความคิดเห็นเกี่ยวกับความสามารถของระบบโดยรวมในแต่ละมอดูล 172
5.21 ความคิดเห็นเกี่ยวกับมอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์ 173
5.22 ความคิดเห็นเกี่ยวกับมอดูลการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์จากภายนอก 173
5.23 ความคิดเห็นเกี่ยวกับมอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์ 174

สารบัญตาราง (ต่อ)

ตารางที่ **หน้า**

5.24 ความคิดเห็นเกี่ยวกับความง่ายในการใช้งานระบบจัดเก็บและสืบค้นเลิร์นนิง อ็อบเจกต์ 174

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญรูป

รูปที่ **หน้า**

2.1 พฤติกรรมการแสวงหาสารสนเทศของเล็กกี้ (Leckie)....13
2.2 พฤติกรรมการแสวงหาสารสนเทศโดยเอลลิส (Ellis)....14
2.3 พฤติกรรมการแสวงหาสารสนเทศของผู้ใช้ของวิลสัน (Wilson)....16
2.4 ข้อมูลสื่อย่อยคือสื่อประสมหรือเนื้อหาที่เป็นองค์ประกอบย่อยๆ ของบทเรียน 28
2.5 วัตถุสารสนเทศเกิดจากการรวมสื่อข้อมูลย่อยหลายชนิดบรรจุเชื่อมโยงร่วมกัน 29
2.6 ระดับเนื้อหาของเลิร์นนิงอ็อบเจกต์ 30
2.7 โครงสร้างและหน่วยข้อมูลย่อยของ LOM 36
2.8 โครงสร้างความสัมพันธ์ของศัพท์ควบคุมประเภทหัวเรื่อง 39
2.9 โครงสร้างความสัมพันธ์ของคำศัพท์ควบคุมประเภทธิซอรัส 40
2.10 รายละเอียดของไฟล์อาร์ดีเอฟที่มีการอ้างถึงข้อมูลยูอาร์ไอ 46
2.11 การแปลงเมทาดาทาแบบระบุองค์ประกอบย่อยของเมทาดาทาที่มีข้อมูลในกลุ่มเดียวกัน....50
2.12 การแปลงองค์ประกอบย่อยเมทาดาทาจากเค้าร่างเมทาดาทาต้นฉบับไปยังเค้าร่างเมทาดาทาเป้าหมาย 51
2.13 วงจรการผสานเมทาดาทา....52
2.14 โครงสร้างความสัมพันธ์ของคลังคำตัวอย่าง....67
2.15 เทคนิคการจัดเรียงลำดับผลการสืบค้น....71
2.16 ปัจจัยที่มีผลต่อการจัดอันดับของระเบียนผลการค้นคืน....72
2.17 กรอบแนวคิดการวิจัย 79
3.1 กรอบแนวคิดการออกแบบชุดเมทาดาทาสำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์ 87
3.2 โครงสร้างเลิร์นนิงอ็อบเจกต์ที่มีการใช้ข้อมูลย่อยร่วมกัน 90
3.3 กรอบแนวคิดการผสานเมทาดาทาและแปลงโครงสร้างเป็นอาร์ดีเอฟ 91
3.4 ความสัมพันธ์ระหว่างบทเรียนและเลิร์นนิงอ็อบเจกต์ 92
3.5 ความสัมพันธ์ระหว่างคลาสบทเรียนและคลาสเลิร์นนิงอ็อบเจกต์....92

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญรูป (ต่อ)

รูปที่ **หน้า**

3.6 การแปลงความสัมพันธ์ของตารางความสัมพันธ์แบบ 1 tom Many แบบที่ 1 93
3.7 การแปลงความสัมพันธ์ของตารางความสัมพันธ์แบบ 1 to Many แบบที่ 2 93
3.8 การกำหนดคุณสมบัติระหว่างคำศัพท์ที่มีความคล้ายคลึงกัน 95
3.9 ลักษณะความขัดแย้งเชิงโครงสร้างสำหรับข้อมูลผู้แต่ง Writer และ Author 96
3.10 ตัวอย่างการแก้ปัญหาความขัดแย้งเชิงโครงสร้าง 96
3.11 ตัวอย่างกฎการผสานเมทาดาทาและแปลงเป็นไฟล์อาร์ดีเอฟ 97
3.12 กรอบแนวคิดขั้นตอนการค้นคืนเลิร์นนิงอ็อบเจกต์ 98
3.13 กรอบแนวคิดการจัดลำดับการนำเสนอผลการสืบค้น 105
5.1 สถาปัตยกรรมโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ 130
5.2 โครงสร้างคลังข้อมูลสำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์ 136
5.3 ฟังก์ชั่นการเพิ่มข้อมูลเลิร์นนิงอ็อบเจกต์ 138
5.4 ฟังก์ชั่นการผสานเมทาดาทาและการแปลงโครงสร้างอาร์ดีเอฟ 139
5.5 คำสั่งแปลงตารางข้อมูลให้เป็นคลาสและพร็อพเพอร์ตี 140
5.6 การคำนวณค่าความคล้ายคลึงของคำศัพท์เมทาดาทา 141
5.7 การกำหนดคุณสมบัติระหว่างคำศัพท์ที่มีความคล้ายคลึงกัน 142
5.8 ลักษณะความขัดแย้งเชิงโครงสร้างสำหรับข้อมูลผู้แต่ง Writer และ Author 143
5.9 ตัวอย่างการแก้ปัญหาความขัดแย้งเชิงโครงสร้าง 143
5.10 ตัวอย่างไฟล์อาร์ดีเอฟหลังจากการผสานเมทาดาทาและแปลงโครงสร้าง 144
5.11 สถาปัตยกรรมโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ 145
5.12 ตัวอย่างกราฟอาร์ดีเอฟเลิร์นนิงอ็อบเจกต์หัวข้อระบบจำนวนเต็ม 147
5.13 โครงสร้างคำศัพท์ทางคณิตศาสตร์เรื่องจำนวนและการคำนวณ 151
5.14 จำนวนคะแนนการเปลี่ยนโหนดโดยเริ่มที่รหัสบทเรียน “ระบบจำนวนเต็ม” 155
5.15 การคิดค่าคะแนนเพื่อกำหนดลำดับการนำเสนอให้กับผลการสืบค้น 159
5.16 ตัวเลือกเพื่อจัดลำดับการนำเสนอผลการสืบค้น 161
5.17 กราฟแสดงค่าน้ำหนัก (p) ที่เหมาะสมกับชุดข้อมูล 163

สารบัญรูป (ต่อ)

รูปที่ **หน้า**

5.18 แผนภูมิสรุปการประเมินความถูกต้องของการสืบค้น 164
5.19 เปรียบเทียบเวลาการสืบค้นระหว่างระบบเจนาและระบบที่พัฒนาขึ้นเอง 166
5.20 เวลาการสืบค้นเมื่อมีการเพิ่มระดับความลึกในการสืบค้น 167
5.21 หน้าหลักของระบบเมื่อเข้าใช้งาน 168
5.22 เมนูย่อยเมื่อเข้าใช้งานคำสั่งการสืบค้นแบบไล่เรียง การนำเข้าข้อมูลและ
การแปลงไฟล์ 168
5.23 การสืบค้นแบบไล่เรียงตามระดับการศึกษาและบทเรียน 169
5.24 ผลการแสดงเลิร์นนิงอ็อบเจกต์แบบไล่เรียงตามชั้นปีการศึกษา 169
5.25 หน้าจอเพิ่มเลิร์นนิงอ็อบเจกต์เข้าในระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ 170
5.26 ผลการสืบค้นด้วยคำสำคัญโดยค้นทุกเขตข้อมูล 170
5.27 ผลการแสดงรายละเอียดของเลิร์นนิงอ็อบเจกต์ที่เลือกจากรายการหลัก 171

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

รูปแบบการศึกษาของประเทศไทยได้มีการพัฒนากระบวนการเรียนการสอนอย่างต่อเนื่อง โดยมุ่งเน้นส่งเสริมให้ผู้สอนสามารถเลือกที่จะพัฒนาสื่อการเรียนรู้ด้วยตนเอง หรือเลือกใช้สื่อประกอบการสอนที่มีความทันสมัย ประกอบด้วยอุปกรณ์และเครื่องมือที่ใช้ประกอบการเรียนการสอนที่รองรับการเปลี่ยนแปลงของสาระความรู้ใหม่ ๆ ที่เกิดขึ้นตลอดเวลา เพื่อให้ผู้เรียนมีความรู้และความสามารถทัดเทียมได้กับประเทศอื่น ๆ ในสังคมโลก อีกทั้งเพื่อให้ประเทศสามารถผลิตแรงงานความรู้ (Knowledge Worker) ที่สามารถใช้ความรู้ต่อยอดและเชื่อมโยงกับความรู้ใหม่จากแหล่งความรู้ในระดับชาติและนานาชาติ เพื่อแลกเปลี่ยนและเชื่อมโยงให้เกิดเป็นองค์ความรู้ที่สามารถปรับใช้ตามการเปลี่ยนแปลงของบริบทสังคมในปัจจุบันและอนาคตได้ (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2554: 9)

แผนการศึกษาแห่งชาติฉบับปรับปรุง (พ.ศ.2552-2559) จึงเน้นแนวนโยบายที่สำคัญคือการเพิ่มโอกาสทางการศึกษาให้ประชาชนทุกคนได้มีโอกาสเข้าถึงบริการการศึกษาและการเรียนรู้ โดยเฉพาะผู้ด้อยโอกาส สอดรับกับพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 หมวด 4 มาตรา 22-25 ที่ส่งเสริมให้ผู้สอนกำหนดรูปแบบการเรียนการสอนที่ผู้สอน และผู้เรียนสามารถเลือกรูปแบบการเรียนรู้ที่เหมาะสมและสอดคล้องกับธรรมชาติการเรียนของแต่ละบุคคล และมีแหล่งเรียนรู้หลากหลายรูปแบบในการเรียนการสอน (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2543: 6)

รัฐบาลจึงส่งเสริมให้ระบบการศึกษามีความยืดหยุ่น ส่งเสริมให้ใช้เทคโนโลยีเพื่อเพิ่มศักยภาพในการเรียนการสอน ทั้งสื่อเทคโนโลยีสารสนเทศเพื่อการศึกษา สื่ออินเทอร์เน็ต และสื่อประสม และส่งเสริมให้มีการนำเทคโนโลยีสารสนเทศเข้ามา ช่วยให้ผู้เรียนสามารถเข้าถึงได้สะดวกและไร้ขีดจำกัดทั้งด้านสถานที่ และเวลา อันจะช่วยส่งเสริมวงจรแห่งการศึกษาตลอดชีวิตให้เป็นวงจรที่ไม่หยุดนิ่งและเพิ่มขยายขอบเขตอย่างไม่สิ้นสุด โดยเน้นและให้ความสำคัญกับผู้เรียนในฐานะที่เป็นศูนย์กลางของการเรียนรู้ มีอิสรภาพที่จะเลือก เนื้อหา รูปแบบ วิธีการในการเรียนรู้ที่เหมาะสมกับตนเอง (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2554: 45)

นอกจากนี้ ยังมุ่งเน้นให้ผู้สอนและผู้เรียนได้มีส่วนร่วมในการกำหนดแนวทางการเรียนอย่างเต็มที่ รัฐบาลจึงได้พัฒนาหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานขึ้นมา โดยให้ที่มีความ

สอดคล้องกับแนวทางของแผนการศึกษา และนำความรู้ในแต่ละท้องที่เข้ามามีส่วนในการจัดเนื้อหาและรูปแบบการเรียนการสอน ให้โอกาสผู้สอนได้พัฒนาเนื้อหารายวิชาตามความเหมาะสมของแต่ละท้องที่ให้สอดคล้องกับลักษณะของผู้เรียนและแนวทางการเรียนการสอน โดยมีกรอบแนวการพัฒนาบทเรียนตามเป้าประสงค์ของหลักสูตรแกนกลางและสาระการเรียนรู้ในแต่ละรายวิชา

ทั้งนี้ การใช้เทคโนโลยีสารสนเทศ ผู้สอนจะมีอิสระในการกำหนดรูปแบบของสื่อประกอบการสอน และกระบวนการเรียนการสอนให้เหมาะสมกับกิจกรรมการสอนและเนื้อหาแต่ละส่วนของบทเรียน นอกจากนี้ยังเปิดโอกาสให้ผู้เรียนได้มีโอกาสเข้ามามีส่วนร่วมในการจัดทำเนื้อหาและสามารถเข้าถึงบทเรียนหรือเนื้อหาการเรียนการสอนได้มากขึ้นโดยไม่จำเป็นจะต้องเข้าเรียนในระบบเสมอไป (สำนักงานคณะกรรมการการศึกษาแห่งชาติ, 2543: 4) ปรับเปลี่ยนให้ทุกสถานที่คือห้องเรียน ผู้สอนสามารถที่จะกำหนดบทเรียน เอกสารเสริม งานมอบหมายและกิจกรรมประกอบการเรียนให้กับกลุ่มผู้เรียน ขณะที่ผู้เรียนสามารถแลกเปลี่ยนข้อมูลและใช้สื่อการสอนในแบบอิเล็กทรอนิกส์โดยผู้เรียนมีส่วนร่วมในการคิด วิเคราะห์หรือแสดงความคิดเห็นได้โดยไม่จำเป็นต้องเรียนในแบบเผชิญหน้า และสามารถเรียนได้ทั้งแบบเวลาจริง (Real-Time) และแบบเรียนรู้ด้วยตนเองผ่านบทเรียนและเนื้อหาการเรียนรู้ที่อาจารย์ผู้สอนได้เตรียมเอาไว้ให้ โดยส่วนประกอบในการเรียนไม่ว่าจะเป็นหนังสือตำราเรียน เอกสารประกอบการบรรยาย หรือแบบฝึกหัด งานมอบหมาย จะถูกจัดทำขึ้นในรูปแบบอิเล็กทรอนิกส์ทั้งหมด อีกทั้งสามารถเชื่อมโยงกับแหล่งข้อมูลจากภายนอกเช่น ระบบห้องสมุดแบบเปิด ระบบคลังข้อมูลและสื่อการเรียนรู้ดิจิทัล เป็นต้น (กรมวิชาการ กระทรวงศึกษาธิการ, 2555)

เทคโนโลยีที่นำมาใช้และมีบทบาทสำคัญต่อการศึกษา ประกอบด้วยเทคโนโลยีสนับสนุนการจัดการศึกษา เช่น ระบบคอมพิวเตอร์ช่วยสอน (Computer Aided Instruction: CAI) ระบบการเรียนอิเล็กทรอนิกส์ (E-Learning) ระบบวิดีโอออนดีมานด์ (Video on Demand) ระบบอินเทอร์เน็ต ตลอดจนระบบการสื่อสารระหว่างบุคคลที่จะช่วยให้ผู้สอนกับผู้เรียนสามารถติดต่อและทำการสอนได้แม้ไม่ได้อยู่ที่เดียวกัน (ชาญชัย พิพัฒน์สันติกุล, 2548: 5)

แม้ว่าการพัฒนาเลิร์นนิงอ็อบเจกต์ (Learning Object) จะไม่ใช่เรื่องใหม่สำหรับผู้สอน อีกทั้งยังมีเครื่องมือในการสร้างเลิร์นนิงอ็อบเจกต์ที่มีประสิทธิภาพหลายตัวให้เลือกใช้ตามต้องการ อย่างไรก็ดี การสร้างเลิร์นนิงอ็อบเจกต์เป็นกระบวนการที่ใช้เวลามาก และมีความยุ่งยากในการจัดทำ ทั้งนี้เนื่องจากผู้สอนต้องหาแหล่งข้อมูลต่าง ๆ อาทิ เนื้อหาประกอบรายวิชา แบบฝึกหัดและส่วนประกอบอื่น (เช่น ภาพประกอบและคำบรรยาย เป็นต้น) เพื่อประกอบขึ้นเป็นบทเรียนด้วยตนเอง โดยเนื้อหาจะต้องมีความเหมาะสม เป็นไปตามหลักสูตรที่กระทรวงศึกษาธิการกำหนด และเหมาะกับระดับการรับรู้ของผู้เรียน (สมวงษ์ แปลงประสพโชค, 2549: 78) อีกทั้งผู้สอนยังต้องศึกษา

ถึงการใช้งานเครื่องมือที่ใช้ในการแปลงรูปแบบเนื้อหาและไฟล์ข้อมูลให้อยู่ในรูปแบบที่สามารถนำไปใช้งานในมาตรฐานที่แตกต่างกันได้ เช่น ภาพเคลื่อนไหวจะต้องมีการบีบอัดให้มีขนาดเล็กโดยไม่เสียคุณภาพมากเกินไป หรือไฟล์เอกสารจะต้องรองรับการแสดงผลได้กับคอมพิวเตอร์ทุกเครื่อง โดยผู้ใช้ไม่จำเป็นต้องติดตั้งโปรแกรมอื่น ๆ เพิ่มเติม เป็นต้น

แนวทางหนึ่งที่จะช่วยให้ผู้สอนได้มาซึ่งเลิร์นนิงอ็อบเจกต์ที่ต้องการโดยไม่ต้องสร้างหรือพัฒนาขึ้นเอง คือ การค้นหาและนำเลิร์นนิงอ็อบเจกต์ที่มีอยู่เดิมกลับมาใช้ซ้ำ เพื่อช่วยลดเวลาและเงินทุนในการจัดทำ โดยอาศัยเสิร์ชเอ็นจินเป็นเครื่องมือในการค้นคืนเลิร์นนิงอ็อบเจกต์ที่กระจัดกระจายอยู่ตามเว็บไซต์ต่าง ๆ หรือค้นคืนจากระบบฐานข้อมูลกลางของหน่วยงานราชการ อย่างไรก็ตาม การค้นคืนเลิร์นนิงอ็อบเจกต์เพื่อมาจัดทำเนื้อหาบทเรียนที่ดีและตรงกับความต้องการของผู้สอนได้ครบถ้วน ยังคงเป็นปัญหาสำหรับครูผู้สอน ผู้วิจัยได้ทำการสำรวจกระบวนการค้นคืนและจัดทำเนื้อหาสำหรับบทเรียนสองภาษาของครูในเขตอำเภอเมือง จังหวัดนครราชสีมา ในช่วงเดือนมกราคม พ.ศ. 2555 พบว่า ปัญหาในการค้นคืนเลิร์นนิงอ็อบเจกต์จากอินเทอร์เน็ตเกิดขึ้นเนื่องจาก (1) ผลลัพธ์มีความซ้ำซ้อนกันโดยมีการนำเนื้อหาเดียวกันมาแสดงไว้ในเว็บหลายเว็บไซต์ เมื่อค้นคืนจะทำให้ผู้ใช้คิดว่าเป็นคนละรายการ แต่ที่จริงแล้วมีเนื้อหาแบบเดียวกัน (2) เนื้อหาของเว็บไซต์ไม่เกี่ยวข้องตามชื่อเรื่อง โดยผู้พัฒนาเว็บมักจะใส่คำค้นที่ผู้ใช้สนใจเอาไว้ แต่เนื้อหาเป็นเรื่องอื่น ๆ ที่ไม่เกี่ยวข้องหรือไม่ตรงประเด็นที่เขียนไว้ที่ชื่อเรื่อง หรือ เป็นข้อมูลแอบอ้างหรือหลอกล่อให้เชื่อมโยงไปยังเว็บอื่น ๆ ที่ไม่เกี่ยวข้อง หรือเพื่อขายสินค้า (3) ข้อมูลไม่เป็นปัจจุบัน และมีข้อผิดพลาด (4) ข้อมูลมีลิขสิทธิ์ไม่สามารถนำมาใช้งานได้ หรือจำเป็นต้องเสียเงินซื้อ และ (5) การประเมินเพื่อคัดเลือกทรัพยากรที่ค้นคืนได้เพื่อนำไปใช้ทำได้ยาก เนื่องจากข้อมูลในการประกอบการตัดสินใจที่ได้จากผลการค้นคืนนั้นไม่เพียงพอ โดยส่วนใหญ่จะแสดงเพียงชื่อเรื่องซึ่งแม้จะมีการนำเนื้อหาบางส่วนที่ตัดจากเนื้อหาหลักของเว็บมาแสดงไว้ แต่ก็ยังไม่เพียงพอที่จะประเมินและคัดเลือกผลการค้นที่คาดว่าจะตรงกับความต้องการได้

นอกจากนี้ จากผลการสำรวจยังพบว่า ผู้สอนส่วนใหญ่ค้นหาข้อมูลจากอินเทอร์เน็ตเป็นหลัก โดยผู้สอนเชื่อว่าจะมีข้อมูลที่เกี่ยวข้องกับที่ตนต้องการ และนำกลับมาใช้ซ้ำ (Reused) ได้โดยไม่ต้องสร้างขึ้นใหม่ทั้งหมด อย่างไรก็ดีเนื่องจากการแสดงข้อมูลและเนื้อหาบนหน้าเว็บไซต์มีข้อมูลหลากหลายรูปแบบปะปนกันไป ทำให้ไม่สามารถประกันคุณภาพและความถูกต้องของเนื้อหา อีกทั้งยังพบว่า ข้อมูลเหล่านี้มีจำนวนไม่น้อยที่เป็นข้อมูลที่ซ้ำซ้อนกัน (Data Replication) เช่น เนื้อหาและองค์ประกอบเดิมแต่อยู่ในโครงสร้างหรือการนำเสนอที่มีองค์ประกอบแตกต่างกัน นอกจากนี้ยังพบว่า เนื้อหาจริงของข้อมูลที่ค้นพบ ผู้พัฒนาหรือเจ้าของข้อมูลจะเก็บไว้ในระบบฐานข้อมูลปิด (Close Data) ที่ไม่ต้องการเปิดเผยข้อมูล หรือกำหนดเงื่อนไขในการเข้าถึงข้อมูลหลายรูปแบบ ทำให้ผู้อื่นไม่สามารถเข้ามาใช้งานได้อย่างอิสระ ทำให้เกิดการปิดกั้น และหา

ผลประโยชน์เพียงผู้เดียว ส่งผลให้ข้อมูลที่เป็นประโยชน์ไม่ถูกเผยแพร่หรือนำไปใช้ในวงกว้างอย่างที่ควรจะเป็น

ในปี 2009 เบอร์เนอรส ลี (Berners-Lee, 2009) ได้นำเสนอแนวคิดการสร้างเครือข่ายเชื่อมโยงข้อมูลแบบเปิดหรือลิงค์โอเพนดาทา (Linked Open Data: LOD) เพื่อส่งเสริมให้ข้อมูลบนเว็บไซต์มีการเปิดให้ใช้งานได้อย่างอิสระและเชื่อมโยงเพื่อเข้าถึงข้อมูลอื่นที่เกี่ยวข้องได้อย่างสะดวก โดยแนวคิดนี้ได้รับการตอบรับและพัฒนาจากกลุ่มผู้เป็นเจ้าของหรือดูแลข้อมูลในแขนงต่าง ๆ ทั้งภาครัฐและเอกชนทำให้ในปัจจุบันกลุ่มข้อมูลที่เชื่อมโยงเป็นกลุ่มเมฆ (Linked Open Data Cloud) มีชุดข้อมูลแบบเปิดเสรี (Open Data) เชื่อมโยงกันเป็นใยแมงมุมจำนวนหลายล้านรายการซึ่งข้อมูลเหล่านี้มีการกำหนดมาตรฐานโครงสร้างข้อมูลภายใต้มาตรฐานอาร์ดีเอฟ (RDF) เพื่อให้ระบบคอมพิวเตอร์สามารถเข้าใจโครงสร้างข้อมูลและรองรับการเชื่อมโยงข้อมูลด้วยเทคโนโลยีเชิงความหมายได้ (Semantic Technology)

โอเพนดาทาเป็นการจัดเก็บข้อมูลโดยใช้โครงสร้างในการอธิบายข้อมูลด้วยมาตรฐานอาร์ดีเอฟ โดยผู้ดูแลเว็บไซต์สามารถกำหนดรายละเอียดข้อมูลและเนื้อหาตามรูปแบบที่ตนเองต้องการ เช่น ชุดเมทาดาทา (Metadata) รูปแบบความสัมพันธ์ของข้อมูล (Relationship) และประเภทข้อมูล (Data Type) ทำให้เว็บไซต์ที่มีข้อมูลเรื่องเดียวกันหรือเกี่ยวข้องกัน สามารถสร้างลิงค์เชื่อมโยงระหว่างกันได้ อีกทั้งยังช่วยให้ระบบคอมพิวเตอร์สามารถเข้าถึงและนำเสนอข้อมูลที่เกี่ยวข้องกันได้อย่างเหมาะสมตามมาตรฐานโครงสร้างข้อมูลที่ออกแบบไว้ ทำให้ผู้ใช้สามารถเข้าถึงข้อมูลที่ตนเองต้องการได้อย่างครบถ้วน แต่เนื่องด้วยปัจจุบันขนาดของข้อมูลในกลุ่มเมฆมีขนาดใหญ่ขึ้นเรื่อยๆ และมีปริมาณระดับล้านรายการ ทำให้เกิดปัญหาในการเข้าถึงข้อมูลจากเว็บไซต์ที่มีการเชื่อมโยงกันอยู่อย่างซับซ้อน

เนื่องด้วยความอิสระของผู้สร้างเนื้อหา และความสะดวกในการเชื่อมโยงข้อมูล ทำให้ชุดข้อมูล (Dataset) มีจำนวนเพิ่มขึ้นอย่างต่อเนื่อง ปัจจุบันมีจำนวนชุดข้อมูลกว่า 9,960 ชุดข้อมูล โดยแต่ละชุดข้อมูลมีจำนวนทริเปิล (Triple) รวมกันกว่า 8 หมื่นล้านทริเปิล (Cyganiak, 2014) ในจำนวนนี้เป็นชุดข้อมูลที่เกี่ยวข้องกับทรัพยากรการเรียนรู้ (Educational Resources) จำนวนหลายหมื่นรายการ จากทุกระดับชั้นการศึกษา และมีเนื้อหาที่จัดเก็บไว้หลากหลายประเภท อาทิ ข้อมูลของสถาบันการศึกษา สื่อการเรียนการสอน บทความวิชาการ หรือผลงานวิจัย สำหรับนักวิชาการและนักวิจัย ซึ่งทำให้เกิดปัญหาในการเข้าค้นคืนข้อมูลเลิร์นนิงอ็อบเจกต์และข้อมูลทางการศึกษาจากลิงค์โอเพนดาทา คือ

1) ได้ผลการค้นเป็นจำนวนมากและชี้ไปยังแหล่งอ้างอิงเดียวกัน หรือมีความซ้ำซ้อนกัน (Zhao and Ichise, 2012) เนื่องจากข้อมูลมีการเชื่อมโยงกันไปมาระหว่างชุดข้อมูล ผู้ใช้ต้องทำการ

เปรียบเทียบข้อมูลเอง ทำให้ต้องใช้เวลามากในการคัดเลือกผลการค้นคืนที่ต้องการ (Noia, Mirizzi, Claudio, Davide and Zanker, M, 2012)

2) เลิร์นนิงอ็อบเจกต์ที่จัดเก็บมีชุดข้อมูล หรือเมทาดาทาที่ใช้การอธิบายข้อมูลแตกต่างกัน ทั้งในส่วนของโครงสร้างของชุดเมทาดาทา (Metadata Schema) จำนวนองค์ประกอบย่อย (Element) ที่ใช้ในการบรรยายและระดับในการบรรยายเลิร์นนิงอ็อบเจกต์ ซึ่งทำให้เกิดปัญหาในการใช้งานข้อมูลร่วมกัน นอกจากนี้ การเชื่อมโยงคลังจัดเก็บเลิร์นนิงอ็อบเจกต์ที่ใช้ชุดมาตรฐานเมทาดาทาที่แตกต่างกัน ยังทำให้เกิดความลำบากในการเลือกเนื้อหาที่สอดคล้องและมีรายละเอียดครบถ้วนตามที่ต้องการ แม้จะมีแนวคิดในการคัดเลือกมาตรฐานกลางในการแปลงองค์ประกอบย่อย หรือแอตทริบิวต์ให้เป็นมาตรฐานเดียวกัน อย่างไรก็ดี ด้วยลักษณะข้อมูลที่แตกต่างกันของแต่ละชุดข้อมูล ทำให้การแปลงข้อมูลยังคงเป็นปัญหา และยังไม่มีโมเดลที่สามารถทำงานได้กับทุกมาตรฐาน (Mile, 2010)

เมทาดาทาที่นิยมใช้เพื่อบรรยายเลิร์นนิงอ็อบเจกต์ ได้แก่ ดับบลินคอร์ (Dublin Core: DC) เลิร์นนิงอ็อบเจกต์เมทาดาทา หรือ ลอม (Learning Object Metadata: LOM) และ สคอร์ม (Sharable Content Object Reference Model: SCORM) โดยดับลินคอร์เป็นมาตรฐานในการบรรยายทรัพยากรสารสนเทศดิจิทัลแบบกว้าง จึงไม่สามารถบรรยายรายละเอียดได้ครอบคลุมทุกด้าน เช่น ระดับความยากง่ายของเลิร์นนิงอ็อบเจกต์ เป็นต้น ส่วนลอม (LOM) พัฒนาขึ้นมาเป็นเมทาดาทาสำหรับบรรยายเลิร์นนิงอ็อบเจกต์ที่จัดเก็บในระบบการเรียนอิเล็กทรอนิกส์ (E-Learning) จึงมีความเฉพาะเจาะจงมากกว่าดับลินคอร์ และได้มีการนำไปปรับใช้ในหลายประเทศ เช่น ยูเค ลอม คอร์ (UK LOM Core) ใช้ในอังกฤษ ทีดับเบิ้ลยูลอม (TWLOM) ใช้ในไต้หวัน (Taiwanese LOM) แอนส์ลอม (ANZ-LOM) ใช้ในออสเตรเลียและนิวซีแลนด์ เป็นต้น มาตรฐานตัวสุดท้ายคือสคอร์ม (SCORM) พัฒนาขึ้นมาเพื่อใช้อธิบายเลิร์นนิงอ็อบเจกต์เพื่อตอบสนองวัตถุประสงค์ในการแลกเปลี่ยนเลิร์นนิงอ็อบเจกต์ หรือการสร้างเลิร์นนิงอ็อบเจกต์ใหม่จากเลิร์นนิงอ็อบเจกต์เดิมที่มีอยู่

3) การพัฒนาส่วนต่อประสานกับผู้ใช้ (User Interface) ในการค้นคืนข้อมูลจากกลุ่มข้อมูลแต่ละชุดยังคงเข้าใจยาก ซับซ้อน จึงไม่เอื้อต่อการเรียนรู้ของผู้ใช้ทั่วไป แต่เหมาะสำหรับผู้พัฒนาระบบในการจัดการข้อมูลเมทาดาทาเท่านั้น และยังต้องมีการพัฒนาระบบสำหรับแปลงคำค้นให้อยู่ในรูปแบบประโยคคำถามด้วยภาษาสปาร์เคิล (SPARQL) ซึ่งรูปคำถามจะต้องปรับเปลี่ยนไปตามเมทาดาทาที่ใช้ในการจัดเก็บข้อมูล ทำให้ให้เกิดปัญหาเมื่อต้องการเข้าถึงชุดข้อมูลจำนวนมากพร้อมกัน (Hagedorn, 2013)

ในช่วงปี ค.ศ. 2009 จนถึงปัจจุบันมีงานวิจัยหลายชิ้นที่มุ่งเน้นการหาวิธีการเข้าถึงและใช้ประโยชน์จากข้อมูลที่อยู่ภายในเครือข่ายกลุ่มข้อมูลเชื่อมโยงแบบเปิด (Linked Open Data: LOD) ปัญหาจากจำนวนรายการข้อมูลที่มีปริมาณมาก อีกทั้งรายการข้อมูลแต่ละรายการมีเมทาดาทาทั้งที่

เหมือนกันและแตกต่างกัน เจา (Zhao and Ichise, 2012: 56) จึงได้ทำการศึกษาเปรียบเทียบโครงสร้างการจัดเก็บข้อมูล หรือโครงสร้างของเมทาดาทา (Metadata Schema) โดยใช้กระบวนการวิเคราะห์แนวคิดที่เหมือนกันและแตกต่างกันระหว่างหลายออนโทโลยี หรือการผสานออนโทโลยี (Ontology Alignment) เพื่อใช้กำหนดโครงสร้างการจัดเก็บข้อมูลที่เหมาะสม รวมถึงนำเสนอวิธีการเปรียบเทียบคลาส (Class) คุณสมบัติ (Properties) และความสัมพันธ์ “เหมือนกันกับ” (SameAs) โดยผลลัพธ์จะได้โครงสร้างสำหรับจัดเก็บข้อมูลที่เหมาะสม ไม่มีความซ้ำซ้อนของคุณสมบัติ และลดความซ้ำซ้อนของรายการข้อมูลได้ ซึ่งมีความคล้ายคลึงกับงานวิจัยของโนอาย (Noia, 2012) ที่นำเสนอแนวทางการเปรียบเทียบความคล้ายคลึงของรายการข้อมูลโดยใช้การวิเคราะห์คุณสมบัติ (Properties) ของโครงสร้างเอกสาร และทำการเปรียบเทียบเอกสารรายการอื่นว่าเป็นรายการเดียวกันหรือไม่ ร่วมกับเทคนิคปริภูมิเวกเตอร์ (Vector Space) เพื่อทำการให้ค่าน้ำหนักของคุณสมบัติที่มีความสำคัญ และสามารถบ่งบอกได้ว่าเป็นเอกสารที่เปรียบเทียบกันนั้นเป็นรายการเดียวกันหรือไม่ ในขั้นตอนสุดท้ายได้พัฒนาระบบแนะนำสารสนเทศตามความต้องการของผู้ใช้โดยเก็บความสนใจของผู้ใช้มาวิเคราะห์เพื่อหารูปแบบ (Pattern) ความสนใจต่อเนื้อหาเรื่องใด ๆ และนำเสนอผ่านระบบได้อย่างถูกต้อง งานวิจัยที่มุ่งเน้นการลดความซ้ำซ้อนของชุดข้อมูลในงานชิ้นอื่น เช่น งานวิจัยของฮาสซันและคณะ (Hassan, Sean and Curry, 2012) นำเสนอวิธีการเปรียบเทียบ ควบรวมรายการข้อมูลที่มีความคล้ายคลึงกันแต่มาจากชุดข้อมูลหลายแหล่งด้วยวิธีการวิเคราะห์คุณสมบัติของโครงสร้างข้อมูลเช่นเดียวกับงานวิจัยชิ้นอื่น แต่ได้เพิ่มกระบวนการให้ผู้ใช้มีส่วนร่วมในการประเมินความถูกต้องของผลลัพธ์การค้นคืน โดยผู้ใช้สามารถส่งความคิดเห็นเกี่ยวกับความถูกต้องเหมาะสมของรายการข้อมูลแต่ละรายการว่าเหมาะสมกับค้นหัวเรื่องหรือกลุ่มเนื้อหาหรือไม่ เพื่อนำไปปรับปรุงกระบวนการค้นคืนของระบบให้ดียิ่งขึ้น และงานวิจัยของหวัง (Wang, 2013) ได้เสนอแนวคิดการลดความซ้ำซ้อนของข้อมูลโดยใช้เวลาที่รวดเร็วและมีประสิทธิภาพมากขึ้นโดยการผสานออนโทโลยีในระดับของคุณสมบัติ (Properties) และอินสแตนส์ (Instance) โดยกำหนดแนวทางการเปรียบเทียบความคล้ายคลึงของคลาสในหลายมิติ เพื่อวิเคราะห์ความเหมือนและสอดคล้องระหว่างรายการเอกสารได้อย่างแม่นยำมากที่สุด

จากปัญหาด้านการค้นคืนเลิร์นนิงออบเจกต์เพื่อนำไปใช้ในการผลิตสื่อการเรียนการสอนของครูผู้สอน และจากแนวคิดการแลกเปลี่ยนข้อมูลของลิงค์โอเพนดาทา ผู้วิจัยจึงมีแนวคิดในการออกแบบและพัฒนาโมเดลการจัดเก็บและคืนคืนเลิร์นนิงอ็อบเจกต์ที่สามารถจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากกลุ่มข้อมูลขนาดใหญ่ ที่สามารถเชื่อมโยงข้อมูลที่เกี่ยวข้องจากชุดข้อมูลหลายแหล่งได้พร้อมกัน โดยผ่านระบบการค้นคืนเพียงระบบเดียว โดยระบบสามารถนำข้อมูลจากภายนอกที่มีโครงสร้างและชุดคำศัพท์ที่แตกต่างกันมาจัดเก็บตามโครงสร้างข้อมูลที่ได้ออกแบบไว้ได้อย่างอัตโนมัติ นอกจากนี้ ผู้ใช้ยังสามารถจัดลำดับการนำเสนอผลการค้นคืนตามข้อมูลที่ผู้ใช้

ต้องการ ทั้งนี้เพื่อให้ครูผู้สอนสามารถจัดเก็บเลิร์นนิงอ็อบเจกต์ที่สร้างขึ้นเองไว้ในระบบ และสามารถค้นคืนเลิร์นนิงอ็อบเจกต์ที่ระบบจัดเก็บแบบโอเพนดาทาได้อย่างสะดวกและนำไปใช้ได้ง่ายขึ้น โดยผู้ใช้สามารถค้นคืนเลิร์นนิงอ็อบเจ็กต์ได้ทั้ง 3 ระดับ ได้แก่ วัตถุสารสนเทศ วัตถุประยุกต์ และหน่วยบทเรียน รวมถึงการค้นคืนสื่อข้อมูลย่อยที่อยู่ภายในเลิร์นนิงอ็อบเจกต์ได้ ทั้งนี้เพื่อให้สามารถนำกลับไปใช้ผลิตเป็นเลิร์นนิงอ็อบเจ็กต์ใหม่ได้

1.2 วัตถุประสงค์การวิจัย

1.2.1 เพื่อศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตของครูผู้สอนรายวิชาคณิตศาสตร์หลักสูตรสองภาษา ระดับมัธยมศึกษา

1.2.2 เพื่อออกแบบ พัฒนาและประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา

1.3 คำถามนำการวิจัย

1.3.1 ครูผู้สอนรายวิชาคณิตศาสตร์หลักสูตรสองภาษา ระดับมัธยมศึกษามีพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตอย่างไร

1.3.2 โมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทาควรมีองค์ประกอบในการจัดเก็บอย่างไร และมีกระบวนการจัดเก็บและค้นคืนอย่างไร

1.4 ขอบเขตการวิจัย

1.4.1 กลุ่มตัวอย่างที่ใช้ในการศึกษาพฤติกรรมการค้นคืนและการใช้งานเลิร์นนิงอ็อบเจกต์คือครูผู้สอนรายวิชาคณิตศาสตร์หลักสูตรสองภาษา ระดับมัธยมศึกษาชั้นปีที่ 1 ถึง 6 ในสังกัดของรัฐบาลและเอกชน

1.4.2 เลิร์นนิงอ็อบเจกต์ในระบบต้นแบบเป็นเลิร์นนิงอ็อบเจกต์ในรายวิชาคณิตศาสตร์ที่ถูกสร้างขึ้นใหม่โดยครูผู้สอน และรองรับการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์จากฐานข้อมูลเชิงสัมพันธ์ภายนอกในรูปแบบเอสเควแอล (SQL) เพื่อจัดเก็บตามโครงสร้างแบบโอเพนดาทาเพื่อสะดวกต่อการเข้าถึงข้อมูลและแลกเปลี่ยนข้อมูลตามมาตรฐานแบบเปิดเสรี

1.4.3 การค้นคืนเลิร์นนิงอ็อบเจกต์ จะรองรับรายวิชาคณิตศาสตร์ระดับมัธยมศึกษาเท่านั้น

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ทราบพฤติกรรมการค้นคืนและใช้งานเลิร์นนิงอ็อบเจกต์ของครูผู้สอนรายวิชาคณิตศาสตร์ในหลักสูตรสองภาษา

1.5.2 ได้โมเดลในการจัดเก็บเลิร์นนิงอ็อบเจกต์ประกอบด้วยองค์ประกอบเมทาดาทาในการจัดเก็บเลิร์นนิงอ็อบเจกต์ และกระบวนการผสานเมทาดาทาจากแหล่งข้อมูลที่มีโครงสร้างแตกต่างกัน

1.5.3 ได้โมเดลในการค้นคืนเลิร์นนิงอ็อบเจกต์ด้วยวิธีการเชิงความหมายและสามารถกำหนดลำดับผลการค้นคืนตามความต้องการของผู้ใช้

1.6 คำอธิบายศัพท์

1.6.1 ครูผู้สอน (Teacher) หมายถึง ครูผู้มีหน้าที่ในการสร้างและกำหนดเนื้อหาในการเรียนการสอนที่สอดคล้องกับหลักสูตรแกนกลางตามข้อกำหนดของกระทรวงศึกษาธิการ โดยเป็นครูสอนในรายวิชาคณิตศาสตร์ในหลักสูตรสองภาษา ในช่วงชั้นมัธยมศึกษาชั้นปีที่ 1 ถึง 6 จากโรงเรียนในสังกัดของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) และโรงเรียนเอกชน

1.6.2 การเรียนอิเล็กทรอนิกส์ (Electronic Learning) หมายถึง การเรียน การสอนในลักษณะหรือรูปแบบใดรูปแบบหนึ่ง ซึ่งการถ่ายทอดเนื้อหานั้น กระทำผ่านทางสื่ออิเล็กทรอนิกส์ เช่น ซีดีรอม เครือข่ายอินเทอร์เน็ต หรือ ทางสัญญาณโทรทัศน์ หรือ สัญญาณดาวเทียม (Satellite) เป็นต้น โดยเนื้อหาการเรียนการสอนจะนำเสนอโดยอาศัยเทคโนโลยีมัลติมีเดีย และโต้ตอบจากผู้เรียน โดยสามารถกำหนดรูปแบบของการเรียนรู้ด้วยตนเอง และสื่อในการนำเสนอบทเรียนสามารถมีได้หลายรูปแบบ สามารถที่จะอยู่ในรูปของการสอนทางเดียว หรือการสอนแบบมีปฏิสัมพันธ์ได้

1.6.3 เลิร์นนิงอ็อบเจกต์ (Learning Objects) หมายถึง สื่อดิจิทัล ที่สามารถสื่อให้ผู้เรียนสามารถเข้าใจในเนื้อหาที่สอนในวัตถุประสงค์ใดวัตถุประสงค์หนึ่งตามที่ออกแบบไว้ได้ เนื้อหามีความเป็นอิสระและสมบูรณ์ในตัวเอง สามารถใช้งานรวมกับเลิร์นนิงอ็อบเจกต์รายการอื่นเพื่อให้มีขอบเขตเนื้อหาการเรียนรู้ที่กว้างขึ้นหรือมีวิธีการนำเสนอได้หลากหลายมากขึ้น โดยเลิร์นนิงอ็อบเจกต์จะสร้างขึ้นจากการเชื่อมโยงสื่อข้อมูลย่อย (Asset) หลายหลายการเข้าด้วยกัน องค์ประกอบภายในของเลิร์นนิงอ็อบเจกต์จะประกอบไปด้วย วัตถุประสงค์การเรียนรู้ กิจกรรมการเรียนรู้และวิธีการปฏิสัมพันธ์ และการประเมินผลจากการเรียนรู้ โดยรายละเอียดองค์ประกอบของเลิร์นนิงอ็อบเจกต์มีความแตกต่างกันตามขอบเขตที่กำหนด ทั้งนี้สื่อนั้น ๆ จะถูกจัดเก็บโดยระบบการจัดเก็บเลิร์นนิงอ็อบเจกต์ที่ออกแบบมาเฉพาะ และสามารถเรียกใช้ได้ตามความต้องการ อีกทั้งยังมีความสามารถในการใช้ซ้ำ (Reuse) โดยสามารถนำส่วนหนึ่งส่วนใดหรือทั้งหมดของเลิร์นนิงอ็อบเจกต์มาใช้งานหรือรวมเข้ากับสื่อการเรียนรู้อื่น ๆ เพื่อสร้างเป็นเลิร์นนิงอ็อบเจกต์ใหม่ได้

1.6.4 การค้นคืนสารสนเทศ (Information Retrieval) หมายถึงกระบวนการคัดเลือก ควบคุม โครงสร้างสารสนเทศจากการจัดหา เพื่อทำการจัดเก็บสารสนเทศตามมาตรฐานที่กำหนดไว้ เพื่อสามารถเข้าถึงและแสวงหาทรัพยากรสารสนเทศ ซึ่งจะประกอบด้วยกระบวนการค้นหา (Search) และไล่เรียง (Browse) โดยจะครอบคลุมเรื่องราวหรือเนื้อหาใดเนื้อหาหนึ่งจากแหล่งจัดเก็บสารสนเทศ โดยใช้วิธีการ รูปแบบ และเทคนิคในการกำหนดโครงสร้างประโยคคำถาม (Query) และเงื่อนไข (Criteria) เพื่อคัดกรองผลการค้นคืนให้ตรงกับความต้องการมากที่สุด

1.6.5 โอเพนดาทา (Open Data) หมายถึง โครงสร้างข้อมูลที่เป็นมาตรฐานแบบเปิดเสรี เพื่อส่งเสริมการใช้ซ้ำของข้อมูล ลดความซ้ำซ้อน ใช้ประโยชน์สูงสุดข้อมูลที่มีอยู่ และเพิ่มมูลค่าแก่ข้อมูลโดยสามารถสร้างการเชื่อมโยงเพื่อเพิ่มขอบเขตและความสัมพันธ์ของข้อมูลมากยิ่งขึ้น โดยหลักการใช้ มีดังนี้ (1)ใช้ URI เป็นชื่อสำหรับทุกสิ่งเพื่อเป็นเสมือนรหัสประจำตัวของทรัพยากรบนเว็บไม่ว่าจะเป็น หนังสือ คน สถานที่ แนวคิด หรือความสัมพันธ์ (2)ใช้เอชทีทีพี ยูอาร์ไอ (HTTP URI) เพื่อให้การค้นหา URI นั้นสามารถใช้โปรโตคอลเอชทีทีพี (HTTP) ในการหารายละเอียดเกี่ยวกับ URI นั้นโดยไม่มีชุดคำสั่งที่ซับซ้อน (3) กำหนดมาตรฐานการแลกเปลี่ยนและค้นคืนเช่น อาร์ดีเอฟ (RDF) อาวล์ (OWL) และสปาร์เคิล (SPARQL)

1.6.6 พฤติกรรมการค้นคืนข้อมูลของผู้ใช้ (User Searching Behavior) หมายถึง วิธีการ กิจกรรม หรือการแสดงออกของผู้ใช้ที่เกิดขึ้นในกระบวนการค้นคืนข้อมูล ประกอบด้วย (1) กระบวนการก่อนการค้นคืน ได้แก่ การวิเคราะห์ความต้องการของตนเอง การกำหนดแหล่งค้นคืน การเตรียมวิธีการค้นคืน การกำหนดคำค้น (2) กระบวนการระหว่างการค้นคืน ได้แก่ การเลือกใช้เครื่องมือการค้นคืน การออกแบบการค้นคืน และ (3) กระบวนการหลังการค้นคืน ได้แก่ การพิจารณา และตัดสินใจเลือกผลการค้นคืน รวมถึงการทำกระบวนการซ้ำเพื่อค้นคืนใหม่ซึ่งขึ้นอยู่กับระดับความพึงพอใจต่อผลการค้นคืนในแต่ละครั้ง โดยปัจจัยที่มีผลกับพฤติกรรมได้แก่ แหล่งข้อมูลที่ใช้ในการค้นคืน เครื่องมือที่ใช้ในการค้นคืน และผลการค้นคืน

1.6.7 ปัญหาการค้นคืนข้อมูล (Searching Problems) หมายถึง อุปสรรคหรือข้อขัดข้องที่เกิดขึ้นระหว่างกระบวนการค้นคืนข้อมูล ซึ่งอาจมีสาเหตุมาจากกระบวนการค้นคืน แหล่งข้อมูลที่ใช้ในการค้นคืน หรือตัวผู้ที่ทำการค้นคืน

1.6.8 การผสานเมทาดาทา (Metadata Mapping) หมายถึง กระบวนการรวมชุดโครงสร้างเมทาดาทาสองชุดที่มีโครงสร้างที่แตกต่างกัน แต่มีข้อมูลหรือโดเมนสารสนเทศเรื่องเดียวกันหรือเกี่ยวข้องกัน โดยการรวมชุดเมทาดาทานั้นจะต้องพิจารณาถึงรายละเอียดของแหล่งเก็บข้อมูล และบล็อกเมทาดาทา (Metadata Block) ซึ่งมีองค์ประกอบ 3 ชั้น (Layer)ได้แก่ ภาษาอธิบายเค้าร่าง (Schema Definition Language) เค้าร่างเมทาดาทา (Metadata Schema) และเมทาดาทาอินสแตนส์ (Metadata Instance) โดยระดับการผสานเมทาดาทาสามารถแบ่งกระบวนการเป็น 3 ระดับ ได้แก่

(1) การผสานโครงสร้างภาษา (Language Mapping) เป็นการปรับเปลี่ยนโครงสร้างข้อมูลที่ออกแบบให้รองรับจากภาษาคอมพิวเตอร์ที่มีไวยกรณ์ที่แตกต่างกันให้มีโครงสร้างการจัดเก็บและรองรับคำศัพท์และไวยกรณ์เพียงชุดเดียว เพื่อสามารถผสานโครงสร้างข้อมูลจากโครงสร้างหนึ่งไปยังโครงสร้างที่ต้องการได้ (2) การผสานเขตข้อมูลเมทาดาทา (Metadata Element Mapping) เป็นการปรับเปลี่ยนชุดคำศัพท์ของเมทาดาทาในโครงสร้างข้อมูลที่ใช้คำศัพท์ที่แตกต่างกัน แต่มีความเกี่ยวข้องกันหรืออยู่ในโดเมนเดียวกัน ให้สามารถผสานเป็นชุดเมทาดาทาเดียวกันได้ และ (3) การแปลงอินสแตนส์ (Instance Transformation) เป็นกระบวนการปรับเปลี่ยนประเภทข้อมูล (Data Type) และข้อมูลภายในอินแสตนส์ เพื่อผสานให้โครงสร้างข้อมูลมีลักษณะการจัดเก็บแบบเดียวกัน

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 2
ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

การปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้องกับงานวิจัยเรื่องการพัฒนาโมเดลสำหรับจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลแบบโอเพนดาทา เป็นการศึกษาเพื่อขยายความและทำความเข้าใจในเรื่องที่ศึกษาอยู่นี้ให้ลุ่มลึก กว้างขวาง และครอบคลุม อันจะนำไปสู่การพิจารณาถึงแนวคิด ทฤษฎี และแนวทางการทำวิจัยต่อไป โดยเนื้อหาภายในบทนี้จะนำเสนอแนวคิดและทฤษฎีครอบคลุมเรื่องทั้งหมด ดังนี้

2.1 พฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของผู้ใช้บนอินเทอร์เน็ต
- 2.1.1 ความหมายของพฤติกรรมแสวงหาสารสนเทศ
- 2.1.2 โมเดลวิเคราะห์กระบวนและปัจจัยของพฤติกรรมการแสวงหาสารสนเทศ
- 2.1.3 แหล่งข้อมูลในการค้นคืนเลิร์นนิงอ็อบเจกต์

2.2 เลิร์นนิงอ็อบเจกต์
- 2.2.1 ความหมายของเลิร์นนิงอ็อบเจกต์
- 2.2.2 คุณลักษณะของเลิร์นนิงอ็อบเจกต์
- 2.2.3 องค์ประกอบของเลิร์นนิงอ็อบเจกต์

2.3 การจัดเก็บเลิร์นนิงอ็อบเจกต์
- 2.3.1 เมทาดาทา
- 2.3.2 คลังคำศัพท์สำหรับการค้นคืนสารสนเทศ
- 2.3.3 ลิงค์โอเพนดาทา
- 2.3.4 การผสานเมทาดาทา

2.4 การค้นคืนเลิร์นนิงอ็อบเจกต์ตามมาตรฐานโอเพนดาทา
- 2.4.1 ภาษาและเครื่องมือสำหรับค้นคืนข้อมูลแบบโอเพนดาทา
- 2.4.2 การวัดความคล้ายคลึงของผลการค้น
- 2.4.3 การจัดลำดับผลการค้นคืน

2.5 งานวิจัยที่เกี่ยวข้อง
- 2.5.1 สรุปงานวิจัยที่เกี่ยวข้อง

มหาวิทยาลัยเทคโนโลยีสุรนารี

2.5.2 บทสรุปและเปรียบเทียบกระบวนการวิจัย

2.5.3 กรอบแนวคิดการวิจัย

2.1 พฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์จากอินเทอร์เน็ต

งานวิจัยเรื่อง “การพัฒนาโมเดลสำหรับจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา” ได้มุ่งเน้นศึกษากระบวนการในการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ของผู้ใช้ เพื่อพัฒนาโมเดลที่มีประสิทธิภาพ สอดคล้องกับลักษณะของเลิร์นนิงอ็อบเจกต์และพฤติกรรมการค้นคืนของครูซึ่งเป็นกลุ่มเป้าหมายหลักในการใช้งานระบบ ดังนั้น จึงต้องทำการศึกษาพฤติกรรมและความต้องการในการค้นคืนเลิร์นนิงอ็อบเจกต์จากอินเทอร์เน็ตของผู้ใช้ เพื่อนำมาเป็นข้อมูลพื้นฐานในการออกแบบและพัฒนาโมเดลต่อไป โดยจำแนกการศึกษาแนวคิดเกี่ยวกับพฤติกรรมการค้นคืนข้อมูลของผู้ใช้ และแนวคิดด้านเลิร์นนิงอ็อบเจกต์ ดังต่อไปนี้

2.1.1 ความหมายของพฤติกรรมแสวงหาสารสนเทศ

ชัชวาลย์ วงษ์ประเสริฐ (2537) ให้ความหมายของพฤติกรรมการแสวงหาสารสนเทศ หมายถึง กิจกรรมที่บุคคลกระทำเพื่อหาข้อมูลข่าวสารที่จะตอบสนองต่อความต้องการที่ตนเอง ได้ตระหนักและเล็งเห็นว่าสำคัญ ความต้องการที่เล็งเห็นนี้จะแสดงออกมาและนำไปสู่การกระทำ ที่ทำให้ผู้ใช้ได้รับสารสนเทศมาครอบครอง

จอห์น และพอล (John and Paul, 1997) ให้ความหมายว่า พฤติกรรมการแสวงหาสารสนเทศ หมายถึง กิจกรรมหรือปฏิสัมพันธ์ที่บุคคลแสดงออกเพื่อแสวงหาสารสนเทศอย่างมีจุดมุ่งหมาย

วิลสัส (Wilson, 2000) ให้นิยามว่า พฤติกรรมการแสวงหาสารสนเทศ หมายถึง การแสวงหา สารสนเทศอย่างมีวัตถุประสงค์ โดยเป็นผลมาจากความต้องการใดความต้องการหนึ่ง ทั้งนี้ในระหว่างแสวงหาสารสนเทศ บุคคลผู้นั้นต้องปฏิสัมพันธ์กับระบบสารสนเทศ ซึ่งอาจเป็นระบบสารสนเทศโดยมนุษย์ เช่น ห้องสมุด หนังสือพิมพ์ หรือระบบสารสนเทศด้วยคอมพิวเตอร์ เช่น เวิลด์ไวด์เว็บ เป็นต้น

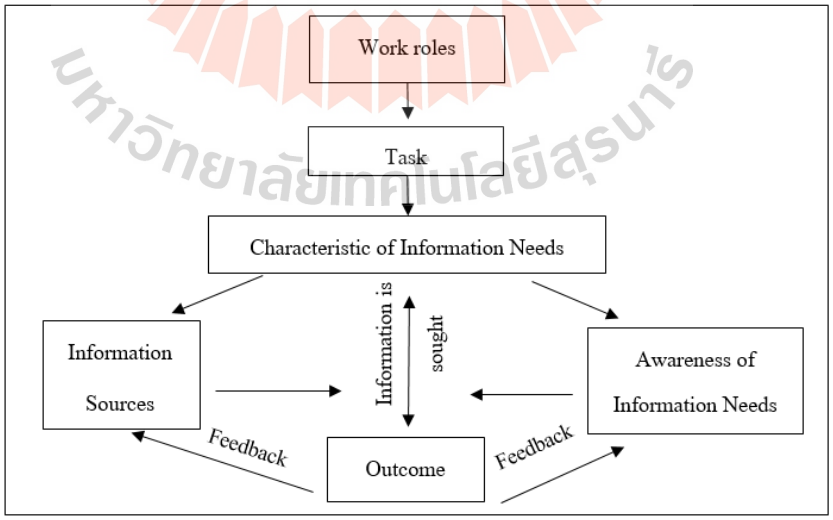
จอห์น และ พอล (John and Paul, 2003) อธิบายว่า พฤติกรรมการแสวงหาสารสนเทศจะเกิดจากความต้องการสารสนเทศ 2 ประการคือ รูปแบบ สารสนเทศที่ต้องการ ว่าผู้ใช้ต้องการค้นคืนสารสนเทศเกี่ยวกับเรื่องใด และมีวัตถุประสงค์ในการค้นคืนสารสนเทศเพื่ออะไร จากนั้นจะส่งผลให้ผู้ใช้แสดงพฤติกรรมการแสวงหาสารสนเทศของตน เพื่อให้ได้มาซึ่งสารสนเทศที่ต้องการ

จากแนวคิดและนิยามข้างต้น ผู้วิจัยจึงขอสรุปว่า พฤติกรรมการแสวงหาสารสนเทศ หมายถึง การกระทำ วิธีการ หรือกิจกรรม ที่แสดงออกมาอย่างมีวัตถุประสงค์ เพื่อแสวงหาข้อมูลสารสนเทศที่ต้องการ จากนั้นสารสนเทศที่ได้จะผ่านการประมวลผล วิเคราะห์ เพื่อนำมาใช้ประโยชน์ หรือสนองความต้องการของตนเอง

2.1.2 แบบจำลองวิเคราะห์กระบวนและปัจจัยของพฤติกรรมการแสวงหาสารสนเทศ

ในการศึกษาพฤติกรรมการแสวงหาสารสนเทศมีการใช้ตัวแบบทฤษฎีแสดงกรอบความคิดเกี่ยวกับพฤติกรรมการแสวงหาสารสนเทศที่นักวิจัยได้พัฒนาขึ้น ดังนี้

เล็กกี้ (Leckie) ปี ค.ศ.1996 ได้เสนอทฤษฎีเกี่ยวกับพฤติกรรมการแสวงหาสารสนเทศของบุคคลที่ประกอบอาชีพวิศวกร แพทย์ และอาจารย์ โดยมีหลักการว่าอาชีพดังกล่าวมีความต้องการสารสนเทศตามภาระงาน ซึ่งจะส่งผลให้เกิดพฤติกรรมการแสวงหาสารสนเทศ โดยมีปัจจัยต่าง ๆ ที่เป็นตัวแปรต่อพฤติกรรมการแสวงหาสารสนเทศ ได้แก่ คุณลักษณะของความต้องการสารสนเทศ การคัดเลือกแหล่งสารสนเทศที่ต้องการ และการตระหนักรู้ในความต้องการสารสนเทศ จากนั้นจะดำเนินการค้นคืนสารสนเทศที่ตนเองต้องการ เพื่อนำไปเป็นข้อมูลในการประกอบอาชีพ และเมื่อได้ข้อมูลเป็นที่น่าพอใจ หรือตรงตามความต้องการแล้ว ก็จะหยุดพฤติกรรมการแสวงหาสารสนเทศนั้น แต่หากผลลัพธ์ที่ได้นั้นยังไม่เป็นที่น่าพอใจ หรือไม่ตรงตามความต้องการ บุคคลนั้นจะเริ่มพฤติกรรมการแสวงหาสารสนเทศใหม่อีกครั้ง ทั้งนี้เล็กกี้ได้แสดงพฤติกรรมดังกล่าว ดังภาพ



รูปที่ 2.1 พฤติกรรมการแสวงหาสารสนเทศของเล็กกี้ (Leckie, 1996: 161)

ปัจจัยที่มีผลต่อพฤติกรรมการแสวงหาสารสนเทศตามแบบจำลองเล็กกี้ ได้แก่

1) อาชีพ (Work roles)

2) ภาระงาน (Tasks)

3) คุณลักษณะของความต้องการสารสนเทศ (Characteristic of Information Needs)

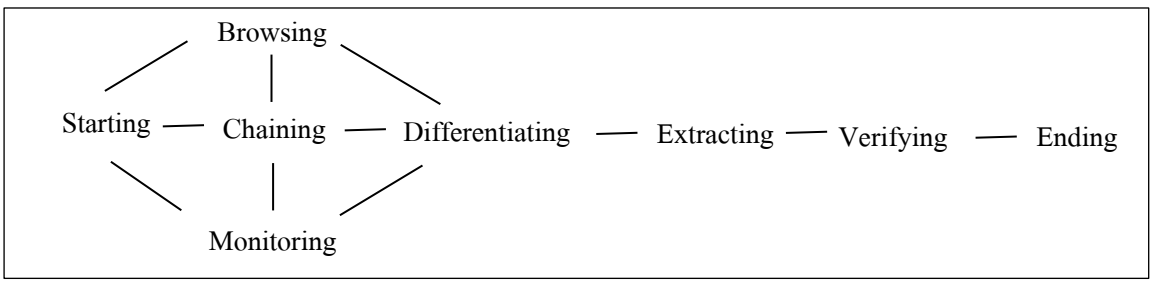
4) การตระหนักรู้ในความต้องการสารสนเทศ (Awareness of Information Needs)

5) การคัดเลือกแหล่งสารสนเทศที่ต้องการ (Source of Information)

6) ผลลัพธ์ (Outcomes)

แบบจำลองหรือโมเดลแสดงให้เห็นถึงพฤติกรรมการแสวงหาสารสนเทศว่า สารสนเทศคือสิ่งที่ต้องแสวงหา และผลที่ได้จากการแสวงหาสารสนเทศคือ ผลลัพธ์ ซึ่งมีส่วนสัมพันธ์กับส่วนอื่น ๆ ของโมเดล ตลอดจนเป็นผลย้อนกลับไปยังแหล่งของข้อมูล การตระหนักรู้และการแสวงหาสารสนเทศ สรุปได้ว่า อาชีพมีส่วนสำคัญในการกำหนดพฤติกรรมการแสวงหาสารสนเทศ นอกจากการได้รับความรู้จากสารสนเทศแล้ว ยังจะต้องมีการแสวงหาสารสนเทศเพิ่มเติม ซึ่งปัจจัยที่เป็นตัวแปรในการปรับเปลี่ยนพฤติกรรมการแสวงหาสารสนเทศที่เกิดขึ้นตลอดเวลา ได้แก่ ปัจจัยด้านสภาพแวดล้อม เป็นต้น เนื่องด้วยอาชีพเป็นตัวกำหนดให้แต่ละบุคคลมีภาระงานที่แตกต่างกัน และ ด้วยภาระงานนั้น ๆ จะส่งผลให้บุคคลมีลักษณะความต้องการสารสนเทศที่แตกต่างกัน

เอลลิสและฮาแกน (Ellis and Haugan, 1997) ได้พัฒนาโมเดลพฤติกรรมการแสวงหาสารสนเทศ ในปี ค.ศ. 1989-1993 โมเดลนี้ประกอบด้วย 6 ขั้นตอน ได้แก่ การเริ่มต้น การเชื่อมโยงการสำรวจ การเลือกดู การแยกแยะ การตรวจตรา และการดึงสารสนเทศออกมา ต่อมาในปี ค.ศ. 1997 Ellis ได้ปรับปรุงโมเดล ดังกล่าวให้สมบูรณ์ โดยเพิ่มอีก 2 ขั้นตอน คือ การตรวจสอบ และการสิ้นสุด รวมเป็น 8 ขั้นตอน ดังรูปที่ 2.2



รูปที่ 2.2 พฤติกรรมการแสวงหาสารสนเทศโดยเอลลิส (Ellis 1997: 391)

จากรูปที่ 2.2 โมเดลพฤติกรรมการแสวงหาสารสนเทศของเอลลิส ประกอบด้วย

1) การเริ่มต้น (Starting) เป็นการเริ่มต้นกระบวนการแสวงหาสารสนเทศ โดยอาจเป็นการทำงานชิ้นใหม่หรือสนใจศึกษาหาความรู้ในเรื่องใหม่ จากการสอบถามเพื่อนร่วมงาน ผู้รู้ หรือการอ่านตำราพื้นฐานเรื่องนั้น ๆ

2) การเชื่อมโยง (Chaining) เป็นการเชื่อมโยงสารสนเทศจากการอ้างอิง หรือบรรณานุกรม อาจเป็นการเชื่อมโยงย้อนหลัง คือเชื่อมโยงจาก รายการอ้างอิงหรือบรรณานุกรมในเอกสารที่มีอยู่ หรือการเชื่อมโยง ข้างหน้า คือเชื่อมโยงว่ามีเอกสารใดอ้างถึงเอกสารที่มีอยู่บ้าง

3) การดูแบบไล่เรียง (Browsing) เป็นการค้นหาโดยมีเรื่องที่ต้องการหรือ สนใจอย่างกว้าง ๆ จึงต้องสำรวจสารสนเทศในสาขาวิชาที่สนใจ เพื่อ เลือกดูหนังสือ บทความที่อาจตรงกับความสนใจเฉพาะได้ เป็นการ เลือกดูอย่างผ่าน ๆ ต่างจากการค้นแบบเฉพาะเจาะจงที่ค้นจากหัวข้อ หรือชื่อเรื่องโดยตรง

4) การแยกแยะ (Differentiating) เป็นการแยกแยะสารสนเทศที่แสวงหาได้ โดยใช้เกณฑ์ต่าง ๆ เช่น ชื่อผู้แต่ง ชื่อวารสาร เป็นเกณฑ์เพื่อกรอง สารสนเทศที่แสวงหา

5) การตรวจตรา (Monitoring) เป็นการตรวจตราวรรณกรรมหรือ สารสนเทศใหม่ ในสาขาวิชาหรือแวดวงวิชาการที่สนใจและคุ้นเคย เช่น การติดตามจากรายชื่อหนังสือใหม่ของสำนักพิมพ์ในสาขาวิชานั้น ๆ การติดต่อสื่อสารกับนักวิจัย

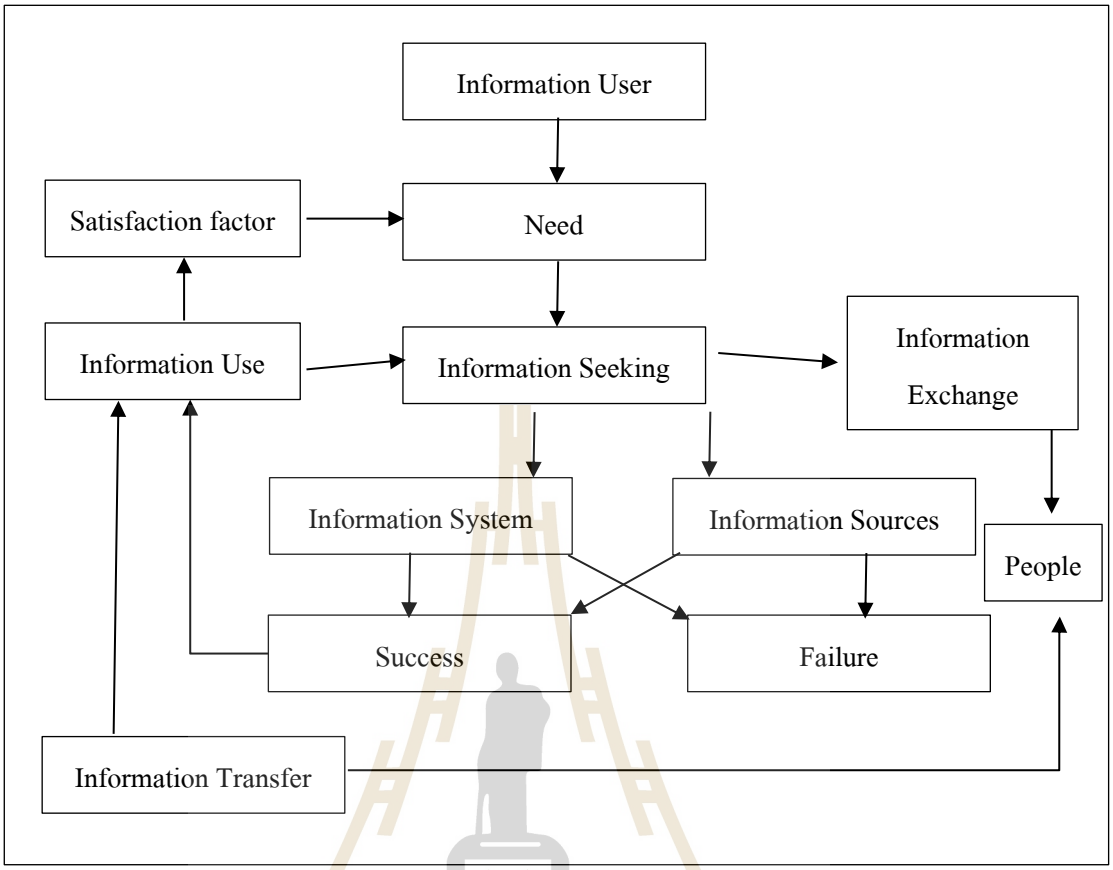
6) การดึงสารสนเทศออกมา (Extracting) เป็นการดึงสารสนเทศที่ต้องการ จากรายงานการวิจัย บทความ วารสาร หนังสือ ฐานข้อมูลดัชนี เอกสาร ประกอบการประชุม ซึ่งเป็นสารสนเทศเพียงบางส่วนในเอกสารที่สามารถนำไปใช้ได้ทันที เช่น ผลการศึกษา ผลวิจัย ผู้แสวงหา สารสนเทศจำเป็นต้องระบุเอกสารที่มีสารสนเทศที่ต้องการก่อน จึงจะ สามารถดึงสารสนเทศออกมาได้

7) การตรวจสอบ (Verifying) เป็นการตรวจสอบความถูกต้องของสารสนเทศที่ได้รับ

8) การสิ้นสุด (Ending) เป็นการแสวงหาสารสนเทศขั้นสุดท้ายเพื่อเก็บรวบรวมสารสนเทศที่แสวงหาได้ทั้งหมดเข้าด้วยกัน เพื่อให้แน่ใจว่าได้สารสนเทศที่ต้องการแล้ว

ทฤษฎีพฤติกรรมการแสวงหาสารสนเทศของเอลลิสเป็นโมเดลพื้นฐานของ การวิจัยเชิงทดลอง และมีการนำไปใช้ในการศึกษาหลายสาขา รวมทั้งผู้ใช้หลายกลุ่ม

Wilson (1999) ได้ศึกษาพฤติกรรมการแสวงหาสารสนเทศตั้งแต่ปี ค.ศ. 1981-1999 อีกทั้งได้พัฒนาโมเดลพฤติกรรมการแสวงหาสารสนเทศหลายโมเดล โดยในการวิจัยการ แสวงหาสารสนเทศ เมื่อปี ค.ศ.1999 วิลสันได้เสนอโมเดลพฤติกรรมการแสวงหาสารสนเทศ ซึ่งประกอบด้วย 12 องค์ประกอบ เริ่มจากผู้ใช้และความต้องการสารสนเทศของผู้ใช้ ที่นำไปสู่พฤติกรรมการแสวงหาสารสนเทศที่แตกต่างกัน และในปี 1999 วิลสันได้พัฒนาโมเดลพฤติกรรม การแสวงหาสารสนเทศที่ครอบคลุมและเป็นกระบวนการ ดังรูปที่ 2.3



รูปที่ 2.3 พฤติกรรมการแสวงหาสารสนเทศของผู้ใช้ของวิลสัน (Wilson, 1999: 251)

จากรูปที่ 2.3 เมื่อผู้ใช้ต้องการสารสนเทศ ผู้ใช้จะแสดงพฤติกรรมการแสวงหาสารสนเทศอย่างเป็นกระบวนการ ดังนี้

1) พฤติกรรมสารสนเทศของผู้ใช้มีจุดผลักดันมาจากความต้องการสารสนเทศ เรียกว่า ความต้องการ พฤติกรรมสารสนเทศครอบคลุม พฤติกรรมหรือกิจกรรมที่เกี่ยวข้อง 2 ด้าน คือ พฤติกรรมการแสวงหาสารสนเทศ และพฤติกรรมการใช้สารสนเทศ

2) ผู้ใช้จะแลกเปลี่ยนสารสนเทศกับบุคคลรอบข้างก่อนเสมอ ซึ่งอาจใช้วิธีการพูดคุยแลกเปลี่ยนความคิดเห็นกับบุคคลต่าง ๆ ไม่ว่าจะเป็นญาติ พี่น้อง เพื่อนร่วมงาน ผู้บังคับบัญชา

3) หากยังไม่ได้รับสารสนเทศที่เพียงพอ ผู้ใช้จะเลือกแสวงหาสารสนเทศ จากระบบสารสนเทศภายใน ซึ่งเป็นระบบสารสนเทศที่ใกล้ตัวที่สุด หากได้รับสารสนเทศตามความต้องการแล้วหมายความว่าการแสวงหาสารสนเทศครั้งนั้นประสบความสำเร็จ แต่ถ้าไม่ได้รับสารสนเทศที่ต้องการ ถือว่าการแสวงหาสารสนเทศครั้งนั้นล้มเหลว

4) เมื่อการแสวงหาสารสนเทศจากระบบภายในล้มเหลว หรือไม่พอใจกับ สารสนเทศที่ได้รับ ผู้ใช้ก็จะเปลี่ยนไปใช้ระบบสารสนเทศอื่น ๆ ที่อยู่ภายนอกแทน

5) หากไม่สามารถแสวงหาสารสนเทศที่ตอบสนองความต้องการ หรือหาไม่พบ อาจต้องย้อนกลับไปค้นหาสารสนเทศอีกครั้ง

6) เมื่อการแสวงหาสารสนเทศประสบความสำเร็จ ผู้ใช้จะนำสารสนเทศที่ ได้รับไปใช้ตามวัตถุประสงค์ นับว่าความต้องการสารสนเทศในครั้งนั้น เป็นที่พึงพอใจของผู้ใช้ ไม่ว่าจะเป็นสารสนเทศที่ได้จากระบบ บริการ หรือแหล่งสารสนเทศก็ตาม ซึ่งสารสนเทศที่ได้รับนี้อาจถ่ายทอดไปยัง บุคคลอื่น ๆ ที่เกี่ยวข้องต่อไป

จากการศึกษาแนวคิดด้านพฤติกรรมการแสวงหาสารสนเทศ การวิจัยครั้งนี้ ได้ประยุกต์ใช้ทั้ง 3 โมเดลในการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน โดยใช้โมเดลการแสวงหาสารสนเทศของวิลสัน (Wilson) ที่แสดงให้เห็นถึงบริบทหรือองค์ประกอบที่เกี่ยวข้องทั้งหมดของการแสวงหาสารสนเทศ ประยุกต์ใช้แนวคิดของเอลลิส (Ellis, 1997) ในการอธิบายกระบวนการของการแสวงหาสารสนเทศ และใช้โมเดลของเล็กกี้ (Leckie, 1996) ในการกำหนดปัจจัยที่เกี่ยวข้องกับการค้นคืนสารสนเทศ โดยโมเดลของเล็กกี้ ได้ระบุไว้อย่างชัดเจนว่าอาชีพ และภาระงานเป็นปัจจัยสำคัญที่มีผลต่อพฤติกรรมการแสวงหาสารสนเทศ ซึ่งสอดคล้องกับงานวิจัยนี้ที่เน้นไปที่ครูคณิตศาสตร์ที่สอนในหลักสูตร 2 ภาษา

2.1.3 แหล่งข้อมูลในการค้นคืนเลิร์นนิงอ็อบเจกต์

2.1.3.1 แหล่งข้อมูลทางอินเทอร์เน็ต

สุมิตรา จิระวุฒินันท์ (2543) กล่าวว่า ความเจริญก้าวหน้าของเทคโนโลยีสารสนเทศ ก่อให้เกิดการพัฒนาการจัดระบบข้อมูลข่าวสารเพื่อให้สอดคล้องกับความต้องการรับรู้ข่าวสารและสารสนเทศของมนุษย์มากขึ้นและรวดเร็วขึ้น เทคโนโลยีสารสนเทศส่งผลต่อรูปแบบการเข้าถึงสารสนเทศและข้อมูลข่าวสารของผู้ใช้โดยตรง เนื่องด้วยความรู้และวิทยาการใหม่ ๆ เกิดขึ้นตลอดเวลาทำให้ความรู้เหล่านั้นไม่สามารถนำเสนอผ่านการเรียนรู้จากหนังสือ วารสารได้อย่างทันท่วงที จึงเกิดช่องทางการนำเสนอข่าวสารและสารสนเทศผ่านระบบช่องทางเครือข่ายอินเทอร์เน็ต ด้วยความสะดวกในการเข้าถึง ความรวดเร็ว ความใหม่ของสารสนเทศที่นำเสนอผ่านระบบอินเทอร์เน็ต ทำให้อินเทอร์เน็ตเป็นช่องทางในการเข้าถึงสารสนเทศที่ได้รับความนิยมมากขึ้นเรื่อย ๆ

ปัจจุบันแหล่งสารสนเทศทางอินเทอร์เน็ตที่ใช้สำหรับค้นคืนเลิร์นนิงอ็อบเจกต์ทั้งในและต่างประเทศมีเป็นจำนวนมาก ผู้วิจัยได้ทำการสำรวจข้อมูลจากกลุ่มตัวอย่าง จำนวน

195 คนโดยวิธีการแจกแบบสอบถามและสัมภาษณ์ เมื่อช่วงปี พ.ศ. 2556 เพื่อทราบแหล่งข้อมูลบนอินเทอร์เน็ตที่ครูผู้สอนนิยมเข้าใช้ สรุปโดยจำแนกเป็นแหล่งข้อมูลแบบปิด และแหล่งข้อมูลแบบเปิดเสรีได้ ดังนี้

1) แหล่งข้อมูลบนระบบปิด

แหล่งข้อมูลระบบปิดหมายถึง แหล่งจัดเก็บสื่อการเรียนการสอน และเลิร์นนิงอ็อบเจกต์ และส่วนประกอบย่อยที่ใช้ในการสร้างสื่อการเรียนการสอน โดยมีการกำหนดสิทธิ์การเข้าใช้งานของผู้ใช้ที่ได้รับอนุญาต รวมถึงมีผู้ดูแลระบบเพื่อตรวจสอบการเข้าถึงสารสนเทศภายในระบบ ทั้งนี้ แหล่งข้อมูลระบบปิดถูกพัฒนาขึ้นเพื่อจัดเก็บข้อมูลสารสนเทศหรือเอกสารที่ต้องการความปลอดภัยและตรวจสอบการใช้งานอย่างเข้มงวด โดยข้อมูลและสารสนเทศนั้นที่มีมูลค่าทางการค้า หรือเป็นข้อมูลที่มีการแสดงสิทธิ์ความเป็นเจ้าของและต้องได้รับอนุญาตก่อนการใช้งาน

จากข้อมูลที่ผู้วิจัยได้ทำการสำรวจแหล่งข้อมูลที่ครูผู้สอนนิยมใช้ที่เป็นแหล่งข้อมูลในการค้นคืนเลิร์นนิงอ็อบเจกต์ที่เป็นระบบแบบปิด ได้แก่ ระบบการเรียนอิเล็กทรอนิกส์ของหน่วยงานราชการหรือสถาบันทางการศึกษา และคลังสารสนเทศดิจิทัลส่วนบุคคล โดยระบบดังกล่าวมีรายละเอียดดังต่อไปนี้

1.1) ระบบการเรียนอิเล็กทรอนิกส์ (Electronic Learning) หมายถึง การเรียนเนื้อหาหรือ สารสนเทศสำหรับการสอนหรือการอบรมซึ่งใช้การนำเสนอด้วยอักษรภาพนิ่ง ผสมผสานกับการใช้ภาพเคลื่อนไหว วีดิทัศน์ และเสียงโดยอาศัยเทคโนโลยีเว็บ (Web Technology) ในการถ่ายทอดเนื้อหา รวมทั้งการใช้เทคโนโลยีระบบการจัดการคอร์ส (Course Management System) ในการบริหารจัดการ งานสอนด้านต่าง ๆ อาทิ การจัดให้มีเครื่องมือการสื่อสารต่าง ๆ เช่น ไปรษณีย์อิเล็กทรอนิกส์ (E-mail), กระดานอิเล็กทรอนิกส์ (Web Board) สำหรับตั้ง คำถาม หรือแลกเปลี่ยนแนวคิดระหว่างผู้เรียนด้วยกัน หรือกับวิทยากร การจัดให้มีแบบทดสอบหลังจากเรียนจบเพื่อวัดผลการเรียนรวมทั้งการจัดให้มีระบบบันทึกติดตามตรวจสอบและประเมินผล การเรียน โดยผู้เรียนที่เรียนจากระบบนี้ส่วนใหญ่แล้วจะศึกษาในลักษณะออนไลน์ ซึ่งหมายถึง การทำการเรียนการสอนผ่านเครื่องมือที่มีการเชื่อมต่อกับระบบเครือข่ายคอมพิวเตอร์ตลอดเวลา

ระบบการเรียนอิเล็กทรอนิกส์ที่ได้รับความนิยมจากข้อมูลการสำรวจได้แก่ มูเดิล (Moodle) ซึ่งเป็นระบบที่สามารถจัดการข้อมูลสารสนเทศของเนื้อการเรียนการสอนและข้อมูลของผู้ใช้และผู้สอนได้อย่างครบถ้วนและที่สำคัญคือสามารถนำมาใช้โดยไม่มีค่าใช้จ่ายอีกด้วย ทั้งนี้มูเดิลถูกนำไปใช้ในหน่วยงานและสถาบันการศึกษาอย่างแพร่หลาย รวมถึงโรงเรียน

ต่าง ๆ ก็นิยมใช้ระบบมูเดิลในการจัดการสื่อการเรียนการสอน เลิร์นนิงอ็อบเจกต์และข้อมูลสารสนเทศที่ผู้สอนต้องการนำเสนอให้กับผู้เรียน และติดตามการเรียนการสอนได้อย่างมีประสิทธิภาพ

1.2) ระบบเว็บไซต์ฐานข้อมูลส่วนบุคคล (Personal Web Database) หมายถึง แหล่งจัดเก็บสารสนเทศที่หน่วยงานหรือสถาบันการศึกษาพัฒนาขึ้นบนเครือข่ายอินเทอร์เน็ต โดยมีวัตถุประสงค์เพื่อจัดเก็บสารสนเทศในรูปแบบต่าง ๆ ของหน่วยงานไว้ร่วมกัน เพื่ออำนวยความสะดวกให้กับผู้ใช้ภายในหน่วยงานสามารถนำข้อมูลสารสนเทศ และเอกสารการสอนไปใช้ประกอบการเรียนการสอน และเผยแพร่ผลงานของตนเองให้กับผู้ใช้ในกลุ่ม หรือในแลกเปลี่ยนกันภายในสถาบันการศึกษาเดียวกัน

จากการสำรวจของผู้วิจัยพบว่าระบบฐานข้อมูลที่พัฒนาโดยเจ้าหน้าที่ภายในหน่วยงานเอง หรือบางสถาบันจะจัดหาระบบจากภายนอกโดยการซื้อหรือว่าจ้างพัฒนาระบบ โดยจะมีการกำหนดฟังก์ชั่นภายในระบบให้ตรงกับความต้องการของผู้ใช้แต่ละหน่วยงานมากที่สุด โดยระบบจะมีฟังก์ชั่นพื้นฐานในการจัดเก็บ ค้นคืนสื่อการเรียนรู้ และเลิร์นนิงอ็อบเจกต์แต่ละรูปแบบได้ รวมถึงการกำหนดสิทธิ์ผู้ใช้ในการเข้าถึง การปรับแก้ หรือการเปิดดูเอกสารต่าง ๆ ได้ตามต้องการ แต่ระบบฐานข้อมูลดังกล่าวมีข้อจำกัดในเรื่องการจัดการระบบในภายหลังเนื่องจากการพัฒนาแบบปิดทำให้ปรับแต่งได้ยาก รวมถึงการถ่ายโอนข้อมูล การปรับปรุงเมทาดาทาของชุดข้อมูล ยังไม่รองรับมาตรฐานแบบสากลเนื่องจากเป็นการพัฒนาตามความต้องการของหน่วยงานแต่ละแหล่ง และมีราคาในการจัดหาและดูแลรักษาสูง

จะเห็นได้ว่าแหล่งข้อมูลแบบปิดนั้นมีข้อดีด้านระบบฟังก์ชั่นที่ครบถ้วน สามารถรองรับสารสนเทศได้หลายรูปแบบ และการจัดการข้อมูลของผู้ใช้ และมีระบบเสริมการทำงานเพื่อรองรับรูปแบบการใช้งานได้หลากหลาย รวมถึงการจัดหาระบบที่มีความสะดวก ทั้งแบบที่ไม่มีค่าใช้จ่าย และแบบมีค่าใช้จ่าย เหมาะกับสถาบันหรือหน่วยงานที่ไม่ต้องการพัฒนาระบบขึ้นเองทั้งหมด โดยเน้นการใช้งานและบริหารข้อมูลในภาพรวม

แต่ข้อจำกัดของแหล่งสารสนเทศแบบปิดได้แก่ มาตรฐานการจัดเก็บข้อมูลที่ออกแบบมาเฉพาะการใช้งานในระบบของตนเอง หรือตามที่หน่วยงานกำหนดเท่านั้น ทำให้เกิดปัญหาในการดูแลบำรุงรักษาข้อมูล เช่น การย้ายระบบไปยังระบบอื่น การนำเข้าและนำออกข้อมูลที่มีรูปแบบจำกัด การเปลี่ยนรุ่นของระบบมีผลต่อรูปแบบการจัดเก็บข้อมูลภายในระบบอาจส่งผลต่อความสมบูรณ์และรายละเอียดของข้อมูลภายในได้ นอกจากข้อจำกัดในเรื่องโครงสร้างข้อมูลภายในระบบแล้ว การเข้าถึงแบบปิดก็เป็นอุปสรรคสำคัญประการหนึ่งที่ทำให้ครูผู้สอนไม่สามารถเข้าถึงเลิร์นนิงอ็อบเจกต์ที่ต้องการได้จากภายนอกหน่วยงานหรือสถาบันที่เก็บข้อมูลเหล่านั้น

จากข้อจำกัดของแหล่งข้อมูลแบบปิดข้างต้น แหล่งข้อมูลที่ครูผู้สอนนิยมใช้เป็นแหล่งค้นคืนสื่อการเรียนรู้ และเลิร์นนิงอ็อบเจกต์เพิ่มเติมจากระบบปิดยังมีบางส่วนที่นิยมใช้แหล่งข้อมูลที่สามารถเข้าถึงได้อย่างเสรี โดยการเข้าถึงผ่านระบบเครือข่ายอินเทอร์เน็ตได้อย่างสะดวก ได้แก่

2) แหล่งข้อมูลบนระบบเปิดเสรี

หมายถึงระบบสารสนเทศบนเว็บไซต์ที่พัฒนาขึ้นและใช้ในการจัดเก็บสื่อการเรียนรู้ เลิร์นนิงอ็อบเจกต์ โดยพัฒนาขึ้นจากหน่วยงานหรือสถาบันการศึกษา และต้องการเผยแพร่สารสนเทศให้กับผู้อื่นนอกเหนือจากบุคลากรภายในองค์กรของตนเอง ให้สามารถเข้าถึงและใช้งานทรัพยากรเหล่านั้นได้อย่างเสรี โดยแหล่งข้อมูลบนระบบเปิดเสรีนั้น จะมีความยืดหยุ่นกว่าระบบปิดในเรื่องการเข้าถึงโดยไม่มีค่าใช้จ่าย ทั้งนี้อาจจะมีการสมัครสมาชิกหรือการยืนยันตนก่อนเข้าใช้งานระบบ จากข้อมูลการสำรวจพบว่าแหล่งข้อมูลบนระบบเปิดเสรีที่ครูผู้สอนนิยมเลือกใช้ ได้แก่

2.1) คลังข้อมูลแบบเปิด (Open Data Repository) หมายถึงแหล่งข้อมูลที่สร้างขึ้นในรูปแบบเว็บแอพลิเคชั่นสำหรับจัดการทรัพยากรสารสนเทศ สื่อการเรียนรู้ และเลิร์นนิงอ็อบเจกต์ผ่านระบบอินเทอร์เน็ต โดยผู้ใช้สามารถเข้าใช้งานระบบได้โดยเสรี และมีผู้ดูแลระบบกำหนดสิทธิ์การใช้งานของผู้ใช้แต่ละบัญชี โดยความเสรีของระบบหมายถึงการเข้าถึง การค้นคืน และการนำข้อมูลไปใช้โดยไม่มีค่าใช้จ่าย แต่ผู้ใช้ทั่วไปไม่สามารถเพิ่มข้อมูล หรือปรับแต่งรายละเอียดของข้อมูลในระบบได้

คลังข้อมูลแบบเปิดนี้ จะถูกพัฒนาโครงสร้างของระบบและข้อมูลตามหน่วยงานหรือสถาบันการศึกษาผู้เป็นเจ้าของสื่อการเรียนรู้ และเลิร์นนิงอ็อบเจกต์ในระบบนั้น ซึ่งการพัฒนาอาจจะใช้โปรแกรมสำเร็จรูปที่มีอยู่แล้วมาปรับแต่งและใช้งาน หรือพัฒนาขึ้นเองตามความต้องการของผู้ใช้ทั้งระบบ ทั้งนี้จากข้อมูลการสำรวจของผู้วิจัย คลังข้อมูลแบบเปิดที่ได้รับความนิยมจากครูผู้สอนเข้าใช้งานจำนวนมาก ได้แก่

เว็บไซต์คลังสื่อการสอนของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ) พัฒนาเว็บไซต์สื่อการสอนการเรียนทางไกล (Distance Learning Thailand: DLIT) ขึ้นเพื่อเป็นแหล่งรวบรวมสื่อการเรียนรู้ และเลิร์นนิงอ็อบเจกต์ที่ตรงตามหลักสูตรเพื่อการศึกษาทางไกล โดยจัดหมวดหมู่และโครงสร้างรายวิชาไว้อย่างชัดเจน โดยครอบคลุมทุกรายวิชาและทุกระดับชั้นปีการศึกษา ตั้งแต่ชั้นประถมศึกษาชั้นปีที่ 1 ถึงระดับชั้นมัธยมศึกษาปีที่ 6 ภายในระบบประกอบด้วยสื่อการสอนในรูปแบบภาพเคลื่อนไหว และเอกสารประกอบการสอนในแต่ละหัวข้ออย่างชัดเจน ผู้เรียนสามารถเข้าเรียนได้โดยผ่านระบบอินเทอร์เน็ต แต่เนื่องจากเป็นเว็บไซต์ที่เน้นให้ผู้เรียนเข้าใช้งาน จึงมีรูปแบบการนำเสนอเพียงภาพเคลื่อนไหวและเอกสารประกอบการเรียนเท่านั้น ไม่มีสื่อรูปแบบอื่น ๆ เช่น ไฟล์เสียง ไฟล์ภาพประกอบ หรือข้อสอบ ให้ผู้สอนนำออกไปใช้ได้มากนัก และประกอบกับการสร้างสื่อการสอนเป็นการสร้างจากครูผู้สอนจากโรงเรียนที่ร่วมโครงการทำให้สื่อการสอนยังมีจำนวนน้อย ทั้งระบบรวม 159 รายการเท่านั้น

เว็บไซต์สื่อการสอนของกระทรวงศึกษาธิการ เป็นเว็บไซต์ที่รวบรวมเสื่อการสอนในทุกระดับชั้นปี และทุกรายวิชาให้เข้าใช้งานได้โดยไม่มีค่าใช้จ่าย แต่สื่อการสอนที่มีจะเป็นไฟล์นำเสนอเท่านั้น ไม่มีสื่อประเภทอื่น ๆ ให้นำไปใช้งาน ระบบเน้นให้ผู้เรียนเข้ามาใช้งานมากกว่าการนำสื่อการสอนออกไปใช้ภายนอกระบบ

เว็บไซต์สื่อประกอบการสอนพัฒนาโดย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท.) เป็นเว็บไซต์ที่มีสื่อการสอนทางด้านวิทยาศาสตร์ คณิตศาสตร์และเทคโนโลยีให้ใช้โดยไม่มีค่าใช้จ่าย ภายในระบบได้กำหนดเนื้อหาโครงสร้างหลักสูตรแกนกลางเพื่อให้ผู้สอนสามารถเลือกเนื้อหาการสอนได้ตรงกับหลักสูตร และเอกสารประกอบการบรรยายหนังสือเรียนตามหลักสูตร และโครงการอบรมต่าง ๆ ซึ่งข้อจำกัดของเว็บไซต์คือสื่อที่มีให้นำไปใช้จะเป็นหนังสือเรียนอิเล็กทรอนิกส์เป็นส่วนใหญ่ ไม่มีเลิร์นนิงอ็อบเจกต์ที่เป็นสื่อการสอนรูปแบบอื่น ๆ ให้ใช้งาน

จากลักษณะการเข้าถึงและค้นคืนข้อมูลจากแหล่งข้อมูลบนอินเทอร์เน็ตของครูผู้สอน จะเห็นได้ว่าแหล่งข้อมูลที่มีอยู่ในปัจจุบันยังไม่มีความเหมาะสมสำหรับการใช้เป็นแหล่งค้นคืนและใช้งานเลิร์นนิงอ็อบเจกต์เท่าที่ควร เนื่องจากข้อจำกัดของแหล่งข้อมูลในหลายประการข้างต้น ผู้วิจัยขอสรุปข้อจำกัดของแหล่งข้อมูลข้างต้น ดังต่อไปนี้

ตารางที่ 2.1 เปรียบเทียบคุณสมบัติของแหล่งข้อมูลแบบระบบปิดและระบบเปิดเสรี

การเปรียบเทียบ	แหล่งข้อมูลบนระบบปิด	แหล่งข้อมูลบนระบบเปิดเสรี
ค่าใช้จ่ายในการใช้งานระบบและสิทธิการใช้ทรัพยากร	มีค่าใช้จ่ายในการจัดหา เข้าใช้งานระบบ ผู้ใช้มีสิทธิ์หลายระดับตามตำแหน่งหน้าที่ในหน่วยงาน	ไม่มีค่าใช้จ่ายในการเข้าใช้งาน สามารถนำทรัพยากรที่มีไปใช้ได้อย่างเสรี
การเข้าถึงทรัพยากร	เฉพาะบุคคลที่ได้รับอนุญาต	ไม่มีข้อจำกัดสำหรับผู้เข้าใช้งาน
การนำเข้า/นำออกข้อมูล	มีรูปแบบการนำเข้าและนำข้อมูลออกตามที่กำหนดไว้	มีรูปแบบการนำเข้าและนำออกหลายแบบตามที่ระบบกำหนดไว้
การเปลี่ยนแปลงโครงสร้างข้อมูลภายใน	ปรับปรุงโครงสร้างข้อมูลไม่ได้ เนื่องจากใช้ระบบฐานข้อมูล	ปรับปรุงโครงสร้างข้อมูลไม่ได้ เนื่องจาก ใช้ระบบฐานข้อมูล

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตารางที่ 2.1 เปรียบเทียบคุณสมบัติของแหล่งข้อมูลแบบระบบปิดและระบบเปิดเสรี (ต่อ)

การเปรียบเทียบ	แหล่งข้อมูลบนระบบปิด	แหล่งข้อมูลบนระบบเปิดเสรี
การรองรับมาตรฐานเมทาดาทาในแบบสากล	ขึ้นอยู่กับการออกแบบโครงสร้างข้อมูลที่จัดเก็บในขั้นตอนเริ่มพัฒนา มักจะปรับเปลี่ยนไม่ได้	ขึ้นอยู่กับการออกแบบโครงสร้างข้อมูลที่จัดเก็บ และปรับเปลี่ยนตามข้อกำหนดของผู้ใช้งานได้ในภายหลัง
การรองรับการแลกเปลี่ยนข้อมูลกับระบบอื่น	รองรับการปรับเปลี่ยนเฉพาะระบบเดียวกันเท่านั้น	สามารถแลกเปลี่ยนข้อมูลกับระบบอื่นได้ แต่ต้องมีโครงสร้างข้อมูลแบบเดียวกัน

จากตารางที่ 2.1 จะเห็นได้ว่าแหล่งข้อมูลบนระบบเปิดนั้นมีความยืดหยุ่นมากกว่าแหล่งข้อมูลบนระบบปิดหลายประเด็น แต่ทั้งนี้แหล่งข้อมูลบนระบบเปิดก็ยังมีข้อจำกัดในเรื่องการแลกเปลี่ยนข้อมูลที่อาจจะไม่รองรับมาตรฐานโครงสร้างข้อมูลของระบบอื่น ๆ และโครงสร้างข้อมูลในอนาคต และการนำเข้าข้อมูลที่ใช้โครงสร้างเมทาดาทาไม่ตรงกับที่ระบบออกแบบไว้จะทำให้เกิดปัญหาในการโอนถ่ายข้อมูล การปรับเปลี่ยนรุ่นของระบบ และการแลกเปลี่ยนข้อมูลกับระบบเปิดเสรีอื่น ๆ ได้

จากข้อจำกัดของแหล่งข้อมูลแบบเปิดเสรีข้างต้น ผู้วิจัยจึงได้หาแนวทางการนำแนวคิดการจัดเก็บเลิร์นนิงอ็อบเจกต์บนระบบแบบเปิดเสรี มาปรับปรุงและพัฒนาให้ระบบเปิดแบบเสรีลดข้อจำกัดดังต่อไปนี้

1) ข้อจำกัดด้านโครงสร้างเมทาดาทาที่ไม่เอื้อต่อการแลกเปลี่ยนข้อมูล หรือไม่สามารถนำเข้าข้อมูลจากระบบภายนอกที่มีโครงสร้างเมทาดาทาไม่ตรงกันได้ การวิจัยครั้งนี้จะทำการวิเคราะห์โครงสร้างเมทาดาทาที่จำเป็นในการจัดเก็บข้อมูล ร่วมกับเทคนิคการผสานเมทาดาทา (Metadata Mapping) เพื่อให้ระบบจัดเก็บเลิร์นนิงอ็อบเจกต์มีโครงสร้างเมทาดาทาแตกต่างกันสามารถแลกเปลี่ยนข้อมูลข้ามระบบได้

2) ข้อจำกัดด้านรูปแบบการจัดเก็บข้อมูลด้วยระบบฐานข้อมูลเชิงสัมพันธ์ ทำให้การค้นคืนมีขอบเขตอยู่เพียงในฐานข้อมูลใดฐานข้อมูลหนึ่ง ทำให้ผลการค้นคืนมีขอบเขตจำกัด ดังนั้นเพื่อช่วยให้แนวคิดแหล่งข้อมูลแบบเปิดเสรีสามารถเชื่อมโยงการค้นคืนไปยังระบบเปิดเสรีอื่น ๆ ได้ งานวิจัยนี้จึงมีแนวคิดโครงสร้างข้อมูลแบบโอเพนดาทามาประยุกต์ใช้เป็นตัวกลางในการเข้าถึงแหล่งข้อมูลแบบเปิดเสรีแทนระบบฐานข้อมูลเชิงสัมพันธ์ ซึ่งจะนำเสนอรายละเอียดในลำดับถัดไป

2.2 เลิร์นนิงอ็อบเจกต์

เลิร์นนิงอ็อบเจกต์หรือ Learning Object: LO มีผู้ศึกษาวิจัยและให้นิยามความหมายไว้หลายรูปแบบ โดยมักจะใช้เรียกสื่อการเรียนการสอนที่มีการออกแบบและพัฒนาขึ้นมาในรูปแบบอิเล็กทรอนิกส์ โดยเลิร์นนิงอ็อบเจกต์จะมีการออกแบบและพัฒนาขึ้นมาจากการประกอบเข้าด้วยกันขององค์ประกอบเนื้อหาการเรียนรู้ย่อย หลาย ๆ ส่วนที่มีหลากหลายรูปแบบ ทั้งภาพ ข้อความ ภาพเคลื่อนไหว เสียงประกอบ ซึ่งอาจได้จากการสร้างขึ้นเอง หรือการนำเนื้อหาย่อยมาจากบทเรียนอื่นและนำมารวบรวมเรียบเรียงและกำหนดวัตถุประสงค์ของเลิร์นนิงอ็อบเจกต์นั้นใหม่ เพื่อมุ่งเน้นผลสัมฤทธิ์ในการเรียนรู้อย่างใดอย่างหนึ่ง ทั้งนี้ ในงานวิจัยบางงานอาจจะมีการกำหนดขอบเขตที่กว้างขวางหรือแตกต่างออกไปบ้าง งานวิจัยได้รวบรวมการให้นิยมของเลิร์นนิงอ็อบเจกต์ไว้ดังนี้

2.2.1 ความหมายของเลิร์นนิงอ็อบเจกต์

ความหมายของเลิร์นนิงอ็อบเจกต์มีการนิยามจากนักวิจัยด้านการศึกษาไว้หลากหลายความหมาย ดังนี้

กิดานันท์ มลิทอง (2548) ได้ให้ความหมายของเลิร์นนิงอ็อบเจกต์ไว้ว่า “เลิร์นนิงอ็อบเจกต์”เป็นหน่วยการสอนขนาดเล็กใช้ในอีเลิร์นนิงที่มีเนื้อหาสาระเป็นอิสระในตัวเอง ภายในเลิร์นนิง อ็อบเจกต์ แต่ละหน่วยจะมีส่วนประกอบของไฟล์ดิจิทัลรูปแบบต่าง ๆ รวมกันอยู่ในหน่วยนั้น ผู้ใช้สามารถนำเลิร์นนิงอ็อบเจกต์ แต่ละหน่วยมาใช้ร่วมกันเพื่อเป็นบทเรียนเรื่องใดเรื่องหนึ่งหรือจะใช้ซ้ำในเรื่องอื่น ๆ ได้ อย่างไม่มีขอบเขตจำกัด

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (สสวท., 2549) ให้คำจำกัดความของเลิร์นนิงอ็อบเจกต์ สรุปได้ว่า เลิร์นนิงอ็อบเจกต์เป็นสื่อดิจิทัลประเภทหนึ่งที่มีลักษณะเฉพาะ คือ เป็นสื่อประสม (Multimedia) ที่ออกแบบมาเพื่อให้นักเรียนบรรลุการเรียนรู้ที่คาดหวังอย่างใดอย่างหนึ่งโดยเฉพาะ โดยแต่ละเรื่องจะนำเสนอแนวคิดหลักและแนวคิดย่อย ๆ ผู้สอนสามารถเลือกใช้เลิร์นนิงอ็อบเจกต์ ผสมผสานกับการจัดการเรียนการสอนแบบอื่น ๆ ได้อย่างหลากหลาย เลิร์นนิงอ็อบเจกต์ มีความแตกต่างกับสื่อดิจิทัลอื่น ๆ ตรงเนื้อหาสาระและกระบวนการเรียนรู้ของนักเรียน เนื่องจากสื่อชนิดนี้ “เน้นการออกแบบกระบวนการเรียนรู้”

สติยา ลังการ์พินธุ์ (2548) กล่าวว่า “ สื่อเลิร์นนิงอ็อบเจกต์ เป็นสื่อที่ออกแบบเพื่อให้นักเรียนเรียนรู้แนวคิดหลักอย่างใดอย่างหนึ่งโดยเฉพาะ สามารถจัดเก็บและค้นหาในระบบดิจิทัลได้สะดวกผู้สอนสามารถนาไปใช้ซ้ำได้ในรูปแบบการเรียนการสอนที่หลากหลายในลักษณะเดียวกันกับตัวต่อเลโก้ ที่สามารถใช้ประกอบเป็นรูปร่างต่าง ๆ และสามารถแยกชิ้นส่วนแล้วนาตัวต่อชิ้นเดิมไปสร้างเป็นรูปร่างขึ้นมาใหม่ได้”

วัฒนา เฉียงเหนือ (2549) ได้ให้ความหมายของเลิร์นนิงอ็อบเจกต์ สรุปได้ว่า เลิร์นนิงอ็อบเจกต์ เป็นสื่อดิจิทัลที่มีลักษณะเฉพาะ คือ เป็นสื่อประสม ที่ออกแบบเพื่อให้นักเรียนบรรลุผลการเรียนรู้ที่คาดหวังอย่างใดอย่างหนึ่งโดยเฉพาะ โดยแต่ละเรื่องจะนาเสนอแนวคิดหลักย่อย ๆ ผู้สอนสามารถเลือกใช้เลิร์นนิงอ็อบเจกต์ ผสมผสานกับการเรียนการสอนแบบอื่น ๆ ได้อย่างหลากหลาย

อนุชัย ธีระเรืองไชยศรี (2549) ได้ให้ความหมายของเลิร์นนิงอ็อบเจกต์ว่า เลิร์นนิงอ็อบเจกต์ หมายถึง สื่อดิจิทัลที่ได้รับการออกแบบมาเพื่อใช้สนับสนุนการเรียนรู้และสามารถนำกลับมาใช้ใหม่ได้ หน่วยของเนื้อหาดิจิทัลที่ได้รับการออกแบบตามแนวคิดใหม่ จะลำดับจากหน่วยขนาดใหญ่เป็นหน่วยขนาดเล็กหลายหน่วย หน่วยเนื้อหาแต่ละหน่วยมีเนื้อหาสมบูรณ์ในตัวเอง เป็นอิสระจากกัน หน่วยเนื้อหาแต่ละหน่วยสามารถนำไปใช้ซ้ำได้ในหลายโอกาส หน่วยเนื้อหาแต่ละหน่วยสามารถนำมาเชื่อมโยงกันเป็นหน่วยเนื้อหาขนาดใหญ่ขึ้นตามลำดับ จนเป็นรายวิชาหรือหลักสูตร สามารถกำหนดข้อมูลอธิบายหน่วยเนื้อหาแต่ละหน่วยเพื่ออำนวยความสะดวกในการค้นหา

ใจทิพย์ ณ สงขลา (2550) ได้ให้ความหมายของเลิร์นนิงอ็อบเจกต์ไว้ว่า เลิร์นนิงอ็อบเจกต์ เป็นการออกแบบเนื้อหาสาระในบทเรียนอิเล็กทรอนิกส์ ให้เป็นหน่วยย่อยที่สุด ให้สามารถนำมาผสานเพื่อการใช้ซ้ำและแลกเปลี่ยนในวัตถุประสงค์ทางการเรียนเดียวกัน สามารถให้การชี้แนะนักเรียน สนับสนุนและกระตุ้นต่อการเรียนรู้

ถนอมพร เลาหจรัสแสง (2550) ได้ให้ความหมายของ เลิร์นนิงอ็อบเจกต์โดยนิยามจากมุมมองทางด้านการศึกษาว่า เป็นลักษณะของหน่วยการเรียนการสอนในรูปแบบดิจิทัลที่มีความสมบูรณ์ในตัวเอง ประกอบด้วยวัตถุประสงค์ เนื้อหา ซึ่งนำเสนอแนวคิดเรื่องใดเรื่องหนึ่งหรือหลายเรื่อง แต่จำเป็นต้องออกแบบให้มีการบูรณาการแนวคิดนั้น ๆ เข้าเป็นเรื่องเดียวกัน โดยมีแบบฝึกหัดเชิงโต้ตอบและหรือแบบทดสอบเพื่อวัดผลการเรียนรู้ของนักเรียน รวมทั้งยังมีขนาดเล็กกะทัดรัด ซึ่งหมายถึง เวลาที่นักเรียนใช้การเรียนรู้แต่ละเนื้อหานั้นไม่ควรเกิน 10 -12 นาที โดยที่ยังคงคุณลักษณะที่สำคัญของเลิร์นนิงอ็อบเจกต์ คือ ความสามารถในการนำกลับมาใช้ใหม่

ศยามน อินสะอาด (2550) ให้ความหมายของเลิร์นนิงอ็อบเจกต์ไว้ว่า เลิร์นนิงอ็อบเจกต์เป็นสื่อการสอนดิจิทัลหรือหน่วยการสอนขนาดเล็กที่สามารถนำกลับมาใช้ใหม่โดยการจัดลำดับเนื้อหาใหม่เกิดเป็นบทเรียนเรื่องใหม่ขึ้นโดยมีองค์ประกอบสำคัญ ได้แก่ วัตถุประสงค์การเรียนรู้ หน่วยการเรียนและแบบทดสอบ

คณะกรรมการมาตรฐานเทคโนโลยีทางการศึกษาของ IEEE (The IEEE's Learning Technology Standards Committee : 2003) ซึ่งมีหน่วยงานที่ดูแลการพัฒนาและสนับสนุนมาตรฐานของเทคโนโลยีที่ใช้ในการเรียนการสอน คือ Learning Technology Standard Committee (LTSC)

ซึ่ง LTSC ได้ให้คำจากัดความของเลิร์นนิงอ็อบเจกต์ไว้ว่า “เลิร์นนิงอ็อบเจกต์เป็นองค์ประกอบใด ๆ ทั้งที่อยู่ในรูปแบบดิจิทัลและไม่อยู่ในรูปแบบดิจิทัล ซึ่งสามารถเรียกใช้ หรือนำกลับมาใช้ใหม่ หรืออ้างอิงถึงได้ในระหว่างเทคโนโลยีที่สนับสนุนการเรียนรู้”

จากรายละเอียดที่สรุปจากแนวคิดในงานวิจัยและเอกสารข้อกำหนดมาตรฐานหลายฉบับข้างต้น ผู้วิจัยสามารถสรุปความหมายของเลิร์นนิงอ็อบเจกต์ที่ใช้ในงานวิจัยนี้ได้ดังนี้ เลิร์น-นิงอ็อบเจกต์ เป็นสื่อดิจิทัล ที่สามารถสื่อให้ผู้เรียนสามารถเข้าใจในเนื้อหาที่สอนในวัตถุประสงค์ใดวัตถุประสงค์หนึ่งตามที่ออกแบบไว้ได้ เนื้อหามีความเป็นอิสระและ สมบูรณ์ในตัวเอง สามารถใช้งานรวมกับเลิร์นนิงอ็อบเจกต์รูปแบบอื่นได้ ทั้งนี้สื่อนั้น ๆ จะถูกจัดเก็บโดยระบบการจัดเก็บเลิร์นนิงอ็อบเจกต์ที่ออกแบบมาเฉพาะและสามารถเรียกใช้ได้ตามความต้องการ อีกทั้งยังมีความสามารถในการใช้ซ้ำ (Reuse) โดยสามารถนำส่วนหนึ่งส่วนใดหรือทั้งหมดของเลิร์นนิงอ็อบเจกต์มาใช้งานหรือรวมเข้ากับสื่อการเรียนรู้อื่น ๆ เพื่อสร้างเป็นเลิร์นนิงอ็อบเจกต์ใหม่ได้

2.2.2 คุณลักษณะของเลิร์นนิงอ็อบเจกต์

เลิร์นนิงอ็อบเจกต์มีคุณลักษณะเฉพาะที่สามารถสรุปได้ 6 ข้อ ดังนี้ (อนุชัย ธีระเรืองไชยศรี, 2551)

1) ความสามารถในการนำกลับมาใช้ใหม่ (Reusability) คือ การนำส่วนประกอบของบทเรียน (Object) ย่อย ๆ ซึ่งเป็นส่วนประกอบของเลิร์นนิงอ็อบเจกต์ใด ๆ กลับมาใช้ใหม่ เช่น การนำไฟล์ภาพจากเลิร์นนิงอ็อบเจกต์หนึ่ง กลับมาใช้ใหม่ในเลิร์นนิงอ็อบเจกต์อีกชิ้นหนึ่ง เป็นต้นทั้งนี้การนำกลับมาใช้ใหม่ยังอาจรวมถึง การนำกลับมาใช้ใหม่ของทรัพยากรวัตถุดิบในการสร้างเลิร์นนิงอ็อบเจกต์ เช่น แม่แบบ ปุ่ม เป็นต้น โดยเลิร์นนิงอ็อบเจกต์หนึ่งเรื่องอาจประกอบไปด้วยสื่อการเรียนรู้หลาย ๆ ชิ้นซึ่งจะต้องประเมินถึงความเหมาะสมของสื่อการเรียนรู้ ถ้าสื่อการเรียนรู้มีองค์ประกอบย่อยจำนวนมากจะง่ายต่อการแยกองค์ประกอบนำกลับมาใช้ใหม่ (Reuse) แต่อาจจะขาดความเป็นอันหนึ่งอันเดียวกันของเนื้อหาได้ ได้ แต่หากสื่อการเรียนรู้มีเนื้อหาเฉพาะเจาะจงมากจนส่วนประกอบของสื่อไม่สามารถแยกเป็นส่วนย่อย ๆ ได้มากพอ จะส่งผลให้การนำกลับมาใช้ซ้ำนั้นทำได้ยากตามไปด้วย จะเห็นได้ว่าการพัฒนาเลิร์นนิงอ็อบเจกต์จะประกอบด้วยสื่อการเรียนรู้หลาย ๆ ประเภทโดยผู้จัดทำอาจทำขึ้นเองหรือนำมาจากสื่อการเรียนรู้อื่น ๆ ที่มีเนื้อหาคล้ายคลึงหรือเหมือนกันมาใช้งานก็ได้

2) ความสามารถในการใช้งานร่วมกัน (Share ability) คือ การใช้งานเลิร์นนิงอ็อบเจกต์แม้ว่าเลิร์นนิงอ็อบเจกต์นั้น จะอยู่บนระบบบริหารจัดการการเรียนรู้ (LMS) หรือ ระบบบริหารจัดการเนื้อหา (LCMS) ที่แตกต่างกัน เป็นความสามารถในการนำเลิร์นนิงอ็อบเจกต์ที่มีอยู่ของแต่ละระบบ มาใช้งานร่วมกันได้

3) ความสามารถในการทำงานร่วมกัน (Interoperability) คือ ความสามารถในการเข้าถึงและใช้งานเลิร์นนิงอ็อบเจกต์ แม้ว่าเครื่องมือที่ใช้ในการเข้าถึงเลิร์นนิงอ็อบเจกต์จะมีความแตกต่างกัน

เช่น การเข้าถึงจากคอมพิวเตอร์ โทรศัพท์เคลื่อนที่ หรือ PDA เป็นต้นจะต้องสามารถแสดงผลได้เช่นเดียวกัน

4) ขนาดกะทัดรัด (Bite-sized/Granularity) คือ เวลาที่ผู้เรียนใช้ในการเรียนรู้เนื้อหาจากเลิร์นนิงอ็อบเจกต์แต่ละเลิร์นนิงอ็อบเจกต์นั้น ไม่ควรเกิน 10-12 นาที (Balatsoukas, Morris and O'Brien, 2008) ซึ่งแตกต่างจากการออกแบบคอมพิวเตอร์ช่วยสอน ในสมัยก่อนซึ่งมีงานวิจัยหลายชิ้นที่สนับสนุนว่าค่าเฉลี่ยของเวลาที่ใช้ในการเรียนรู้เนื้อหาที่เหมาะสมของคอมพิวเตอร์ช่วยสอน จะอยู่ที่ประมาณไม่เกิน 25 นาทีต่อการเรียนรู้ของผู้เรียนแต่ละครั้งเนื่องจากบทเรียนมีองค์ประกอบที่ไม่สามารถแยกออกจากกันได้ อีกทั้งยังใส่กิจกรรมต่าง ๆ ไว้รวมกันทั้งหมด ซึ่งเมื่อเริ่มเรียนแล้วต้องดำเนินกิจกรรมให้เสร็จสิ้นจึงใช้เวลานานกว่าเลิร์นนิงอ็อบเจกต์จากแนวคิดปัจจุบัน

5) ความสมบูรณ์ในตนเอง (Self-contained/Integrity) คือ การที่เลิร์นนิงอ็อบเจกต์นั้น ๆ จะมีความสมบูรณ์ในตนเองจะต้องประกอบด้วยวัตถุประสงค์ เนื้อหา แบบฝึกหัดหรือแบบทดสอบ ทั้งนี้ อาจต้องนำหลักการการวิเคราะห์และออกแบบบทเรียนคอมพิวเตอร์ช่วยสอน มาช่วยในการสร้างเลิร์นนิงอ็อบเจกต์

6) เอื้อให้เกิดการเรียนรู้ที่มีความหมาย (Conducive to learning) คือการออกแบบและพัฒนาเลิร์นนิงอ็อบเจกต์ที่ส่งเสริมให้ผู้เรียนเกิดการเรียนรู้ในลักษณะที่สามารถนำไปเชื่อมโยงกับประสบการณ์จริงของผู้เรียนได้ ดังนั้นเลิร์นนิงอ็อบเจกต์ที่สร้างขึ้นจะต้องออกแบบให้สิ่งแวดล้อมการเรียนรู้ของผู้เรียนมีความใกล้เคียงกับความเป็นจริงที่สุดทั้งนี้ เพื่อให้ผู้เรียนสามารถถ่ายโยง (Transfer) ทักษะที่ได้รับจากการใช้เลิร์นนิงอ็อบเจกต์ ดังกล่าวไปใช้ในบริบทอื่น ๆ ต่อไปได้ โดยเลิร์นนิงอ็อบเจกต์ที่สามารถเอื้อต่อการเรียนรู้ในลักษณะดังกล่าวมักได้รับการออกแบบให้อยู่ในรูปแบบของเกมในลักษณะการค้นพบ หรือการสำรวจ เป็นต้น

2.2.3 องค์ประกอบของเลิร์นนิงอ็อบเจกต์

2.2.3.1 จำแนกตามวัตถุประสงค์ของเลิร์นนิงอ็อบเจกต์

เลิร์นนิงอ็อบเจกต์ที่สร้างขึ้นมานั้นอาจมีการกำหนดวัตถุประสงค์การนำไปใช้งานที่แตกต่างกัน แม้ว่าเนื้อหาและรายวิชาที่จะใช้สอนเป็นรายวิชาเดียวกันและเรื่องเดียวกันก็ตาม โดยเนื้อหาภายในจะถูกปรับให้เหมาะกับผู้เรียนตามวัตถุประสงค์ของการเรียน ซึ่งจะส่งผลถึงวิธีการนำเสนอเนื้อหาภายในเลิร์นนิงอ็อบเจกต์แต่ละรายการด้วย เลิร์นนิงอ็อบเจกต์มีแนวทางการแบ่งประเภทของสื่อการเรียนรู้ไว้หลายระดับ ตามตัวแปรที่เกี่ยวข้องตามบริบทของหน่วยงาน เช่น ระดับชั้นปี ประเภทกิจกรรมที่ใช้ประกอบการเรียนการสอน รูปแบบการนำเสนอ เป็นต้น โดยเดวิด เอ ไวรี่ (Wiley, 2000:21) ได้เสนอแนวทางการแบ่งประเภทของเลิร์นนิงอ็อบเจกต์ตามวัตถุประสงค์ไว้ 5 รูปแบบ ได้แก่

1) เลิร์นนิงอ็อบเจกต์พื้นฐาน (Fundamental Learning Object) เป็นเลิร์นนิงอ็อบเจกต์ที่สร้างขึ้นจากสื่อเพียงชนิดเดียวไม่มีการรวมสื่อประเภทอื่นเข้ามาไว้ด้วยกัน มักจะใช้เพื่อเสริมการนำเสนอในการแสดงนิทรรศการหรือผลงานต่าง ๆ เช่น ภาพถ่าย รูปภาพ หรือ เสียงเพลง เสียงพูด เสียงร้องสัตว์ เป็นต้น (Wiley & Nelson, 1998) สื่อการเรียนรู้ประเภทนี้สามารถดูและเข้าใจได้โดยไม่ต้องอธิบายเพิ่มเติมหรือศึกษาส่วนเสริมจากสื่ออื่น ๆ

2) เลิร์นนิงอ็อบเจกต์แบบรวมผสาน (Combine-closed Learning Object) เป็นเลิร์นนิงอ็อบเจกต์ที่มีเลิร์นนิงอ็อบเจกต์จำนวนไม่มากรวมประกอบเป็นเลิร์นนิงอ็อบเจกต์ หนึ่งเรื่อง โดยผู้สร้างได้กำหนดรายละเอียดการใช้งาน ลำดับเนื้อหา แนวทางการนำเสนอไว้เรียบร้อยแล้ว ทั้งนี้องค์ประกอบต่าง ๆ จะเสริมให้ผู้เรียนมีความเข้าใจในเนื้อหามากขึ้นในหลาย ๆ รูปแบบ มีการผสานและเสร็จสมบูรณ์ในตัวเอง โดยเลิร์นนิงอ็อบเจกต์ประเภทนี้จะไม่สามารถนำไปใช้ซ้ำ (Reuse) ในวงกว้างได้ เนื่องจากเนื้อหาองค์ประกอบที่นำมารวมกันได้ผสานเข้าเป็นเนื้อเดียวกันและได้ผลลัพธ์ออกมาเพื่อใช้ในบางวัตถุประสงค์เท่านั้น เช่น ภาพยนตร์ จะประกอบด้วยภาพเคลื่อนไหว เสียงประกอบ ดนตรี คำพูด คำบรรยาย พร้อมทั้งกิจกรรมเสริมภายในเนื้อหา อย่างไรก็ดี เนื้อหาและความซับซ้อนยังมีระดับต่ำ เพื่อเน้นให้เกิดการเรียนรู้และจดจำได้ง่าย

3) เลิร์นนิงอ็อบเจกต์แบบเชื่อมโยง (Combine-Open Learning Object) เป็นเลิร์นนิง อ็อบเจกต์ที่มีขนาดใหญ่ โดยเนื้อหาจะรวบรวมจากการแสดงผลจากเลิร์นนิงอ็อบเจกต์อื่น ๆ ภายในระบบหรือภายนอกระบบที่มีการเชื่อมโยงกันเอาไว้ ทั้งนี้เลิร์นนิงอ็อบเจกต์ประเภทนี้สามารถใช้คุณสมบัติการใช้ซ้ำ (Reuse) ได้ เนื่องจากองค์ประกอบที่รวมกันเป็นสื่อการเรียนรู้นั้นมาจากหลากหลายแหล่ง และถูกเรียกใช้แบบเวลาจริง (Real time) โดยมีการปรับแต่งและเลือกเฉพาะส่วนที่ต้องการมานำเสนอได้ ตัวอย่างเลิร์นนิงอ็อบเจกต์ลักษณะนี้เช่น เว็บไซต์สื่อการเรียนการสอนโดยองค์ประกอบของสื่อการเรียนรู้ประกอบไปด้วย ภาพนิ่ง ภาพเคลื่อนไหว เนื้อหาประกอบ และแหล่งเชื่อมโยงไปยังข้อมูลอื่น ๆ ที่เกี่ยวข้อง ซึ่งเนื้อหาแต่ละส่วนนั้นอาจมาจากแหล่งอื่น ๆ หรือผู้สร้างรวบรวมแล้วนำมาปรับแต่งเพื่อนำเสนอให้เหมาะสมกับเนื้อหาและวัตถุประสงค์ตามต้องการ ทั้งนี้ บทเรียนประเภทนี้สามารถมีแบบฝึกหัด แบบทดสอบ และโจทย์ปัญหาที่ต้องใช้ความรู้มากขึ้นและมีการแลกเปลี่ยนข้อมูลกับชั้นเรียนอื่น ๆ โดยเนื้อหาภายในจะรวมมาจากเลิร์นนิงอ็อบเจกต์พื้นฐาน หรือเลิร์นนิงอ็อบเจกต์แบบรวมผสานหลาย ๆ ชิ้นเพื่อนำมาสร้างเป็นบทเรียนที่มีเนื้อหาครอบคลุมครบถ้วนและประกอบด้วยแบบฝึกหัด แบบทดสอบที่สมบูรณ์ได้

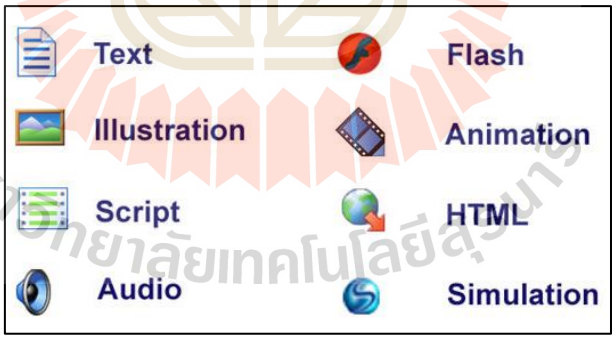
4) เลิร์นนิงอ็อบเจกต์เพื่อการนำเสนอ (Generative Presentation Learning Object) เลิร์นนิงอ็อบเจกต์ประเภทนี้จะรวบรวมสื่อการเรียนรู้อื่น ๆ ในระดับย่อยกว่า เช่น สื่อการเรียนรู้ขั้นพื้นฐาน สื่อการเรียนรู้รวมผสาน หรือสื่อการเรียนรู้แบบเชื่อมโยง โดยการรวมกันจะมีโครงสร้างและวิธีการนำเสนอแบบเฉพาะ เพื่อใช้ในการลำดับความคิดหรือการอ้างอิงภาพรวมของ

เนื้อหาเป็นหลัก เช่น เว็บไซต์สอนพัฒนาโปรแกรม จะมีเนื้อหาทั้งในเว็บไซต์และเชื่อมโยงไปยังแหล่งความรู้ที่มีรายละเอียดมากขึ้นภายในและภายนอกเว็บไซต์ด้วย โดยเนื้อหาการนำเสนอนั้นจะประกอบไปด้วยภาพเคลื่อนไหว ข้อความอธิบาย และตัวอย่าง ซึ่งเนื้อหาแต่ละส่วนจะถูกเรียบเรียงและนำเสนอเป็นทางเลือกเพื่อให้ผู้เรียนสามารถเลือกเรียนตามลำดับ

5) เลิร์นนิงอ็อบเจกต์เพื่อการเรียนการสอน (Generative Instructional Learning Object) เป็นเลิร์นนิงอ็อบเจกต์ที่คล้ายกับเลิร์นนิงอ็อบเจกต์เพื่อการนำเสนอ แต่จะมีรายละเอียดในด้านการทดสอบและประเมินผลตามเนื้อหาประกอบการเรียนรู้ ทั้งนี้ เพื่อสามารถบันทึกผลการเรียนการสอนและสามารถประเมินรูปแบบวิธีการสอน เพื่อนำไปปรับปรุงเนื้อหาการเรียนการสอนให้มีความเหมาะสมยิ่งขึ้น

2.2.3.2 ระดับเนื้อหาในเลิร์นนิงอ็อบเจกต์ (Mallison, 2001, Balatsoukas, Morris, and O'Brien, 2008)

เมื่อวิเคราะห์ภายในเลิร์นนิงอ็อบเจกต์แต่ละรายการแล้วจะพบส่วนประกอบย่อยที่สุดเรียกว่า สื่อข้อมูลย่อย (Media Data Element/Asset) ซึ่งองค์ประกอบเหล่านี้จะมีเนื้อหาหรือการนำเสนอที่สมบูรณ์โดยไม่สามารถแยกย่อยรายละเอียดไปได้อีก หน้าที่ของส่วนประกอบเหล่านี้สามารถนำเสนอเนื้อหาหรือแปลความหมายได้อย่างใดอย่างหนึ่ง และสามารถเปลี่ยนวัตถุประสงค์ได้เมื่อมีการเชื่อมโยงและเรียบเรียงเข้ากับองค์ประกอบชิ้นอื่น ๆ ตัวอย่างองค์ประกอบย่อยเหล่านี้ เช่น รูปภาพประกอบแต่ละภาพ ข้อความ ตัวอักษรที่เรียบเรียงเป็นประโยค โค้ดโปรแกรมแต่ละคำสั่ง ภาพเคลื่อนไหว เสียงเพลง เสียงพูด เป็นต้น ดังรูปที่ 2.4

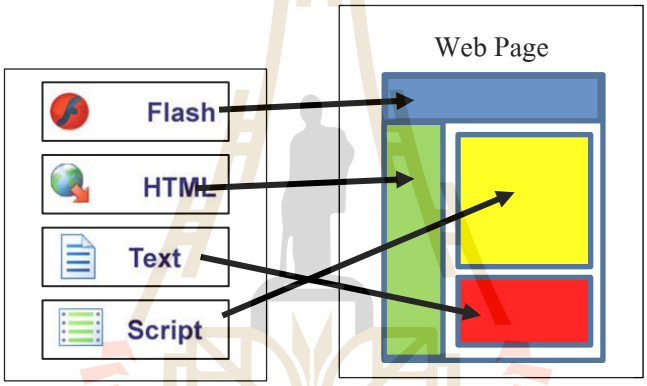


รูปที่ 2.4 ข้อมูลสื่อย่อยคือสื่อประสมหรือเนื้อหาที่เป็นองค์ประกอบย่อย ๆ ของบทเรียน

เมื่อนำสื่อข้อมูลย่อยข้างต้น มาเรียบเรียง เชื่อมโยง และกำหนดวัตถุประสงค์การนำเสนอสื่อข้อมูลย่อยแต่ละส่วน รวมถึงกำหนดการนำเสนอเนื้อหาการเรียนรู้ การมีปฏิสัมพันธ์ระหว่างข้อมูลและผู้ใช้อย่างเหมาะสม รวมถึงกำหนดแนวทางการประเมินผลการเรียนรู้ จะถือว่าสื่อ

นั้นเป็นเลิร์นนิงอ็อบเจกต์ ทั้งนี้ เลิร์นนิงอ็อบเจกต์สามารถแบ่งตามระดับเนื้อหาภายในได้เป็น 3 ระดับ ได้แก่

1) วัตถุสารสนเทศ (Information Object) คือสื่อการเรียนรู้ที่มีการเรียบเรียงขึ้นจากการรวมชิ้นส่วนของสื่อข้อมูลย่อย (Media Data Elements) เข้าด้วยกัน ซึ่งสื่อข้อมูลย่อยแต่ละชิ้นจะถูกกำหนดบทบาทหน้าที่ใหม่ตามลำดับโครงสร้างเนื้อหาของวัตถุสารสนเทศ และวัตถุประสงค์ที่มีขอบเขตมากขึ้นตามเนื้อหาที่กำหนดไว้ โดยวัตถุสารสนเทศจะมีความซับซ้อนมากกว่าสื่อข้อมูลย่อย เนื่องจากมีการเชื่อมโยงเข้ากับสื่อข้อมูลย่อยอื่นตามกรอบวัตถุประสงค์และผลลัพธ์ที่คาดหวังจากการเรียนรู้จากวัตถุสารสนเทศเพียงบางวัตถุประสงค์เท่านั้น แสดงได้ดังรูปที่ 2.5 ตัวอย่างของวัตถุสารสนเทศ เช่น ข้อสอบหนึ่งชุด แบบฝึกหัดหนึ่งชุด เนื้อหาบทเรียน เป็นต้น

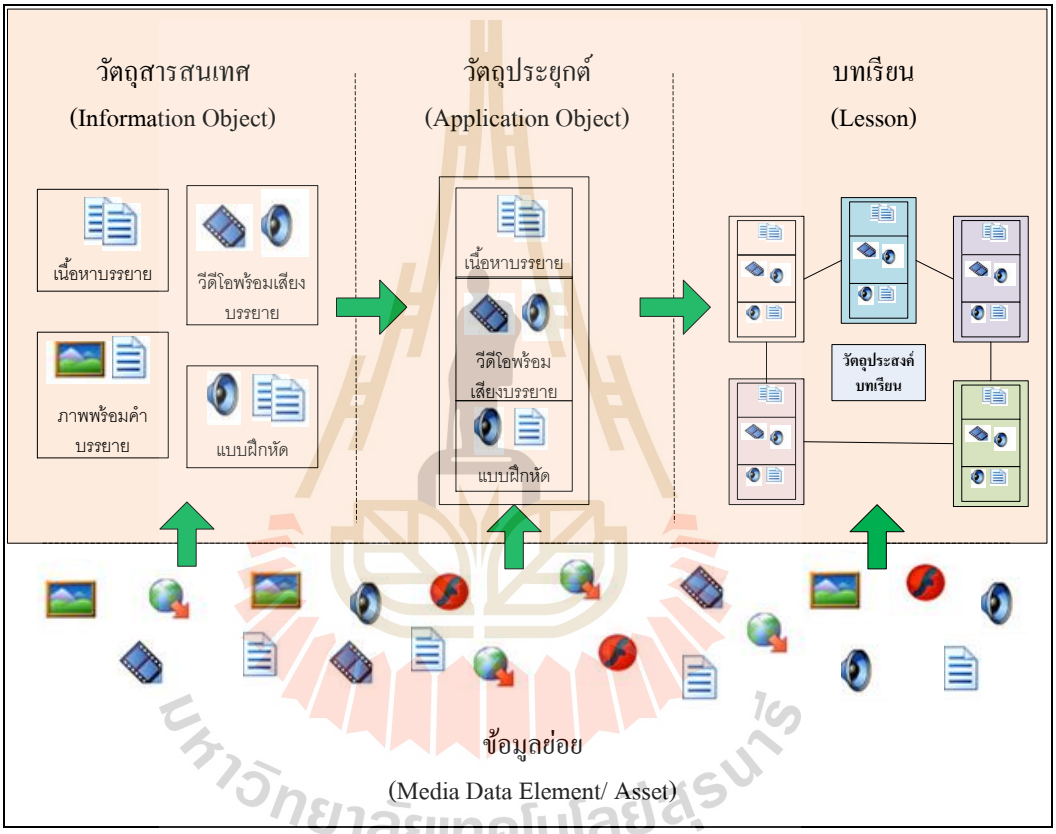


รูปที่ 2.5 วัตถุสารสนเทศเกิดจากการรวมสื่อข้อมูลย่อยหลายชนิดบรรจุเชื่อมโยงร่วมกัน

2) วัตถุประยุกต์ (Application Object) เป็นการเลือกวัตถุสารสนเทศที่กระจัดกระจายมาเชื่อมโยงตามวัตถุประสงค์หลักแต่ละข้อของบทเรียน โดยวัตถุเนื้อหาจะมีความสมบูรณ์ทางด้านเนื้อหาตามวัตถุประสงค์ตามที่กำหนด และมีการลำดับการนำเสนอที่เป็นขั้นตอนชัดเจน ใช้เพื่อบรรลุวัตถุประสงค์การเรียนรู้เพียงข้อใดข้อหนึ่งเท่านั้น ตัวอย่างของเลิร์นนิงอ็อบเจกต์ เช่น ข้อสอบก่อนเรียนหนึ่งชุด เนื้อหาบทเรียนพร้อมแบบฝึกหัด ของเล่นจำลองการชั่ง วัด ตวง วีดีโอสื่อการสอนเรื่องจำนวนเต็มบวก เป็นต้น

3) หน่วยบทเรียน (Lesson Unit or Aggregate Assemblies) คือการเรียบเรียงเชื่อมโยง จัดกลุ่ม วัตถุประยุกต์เข้าด้วยกันตามโครงสร้างรายวิชาในหลักสูตรการเรียนรู้ และวัตถุประสงค์ของบทเรียนแต่ละบทเรียน ซึ่งบทเรียนนั้นอาจจะเรียกว่า กลุ่มวัตถุประยุกต์ (Application Object Chunks) ซึ่งจะมีการกำหนดวัตถุประสงค์และผลสัมฤทธิ์หลักของการเรียนรู้

จากการใช้วัตถุประยุกต์ทุกชิ้นอย่างครบถ้วนและเป็นระบบ โดยมีองค์ประกอบเป็นลำดับการเรียนรู้ชัดเจน เช่น ก่อนเรียนต้องมีการทำแบบทดสอบก่อนเรียน มีเนื้อหาบทเรียนพร้อมตัวอย่างประกอบและแบบฝึกหัดท้ายบท ก่อนจะมีการทดสอบหลังเรียน เป็นต้น โดยแสดงได้ตามรูปที่ 2.6 ซึ่งผู้เรียนต้องใช้วัตถุประยุตก์หลายตัวที่ร้อยเรียงประกอบกันขึ้นเป็นบทเรียนแต่ละบทเรียน เพื่อให้เกิดความรู้รอบด้านตามวัตถุประสงค์หลักของบทเรียนนั้น ตัวอย่างของบทเรียน เช่น บทเรียนคณิตศาสตร์เรื่องจำนวนเต็มบวก เว็บไซต์สอนการเขียนโปรแกรมเชิงวัตถุด้วยภาษาจาวา เป็นต้น



รูปที่ 2.6 ระดับเนื้อหาของเลิร์นนิงอ็อบเจกต์ (Mallison, 2001)

จากรูปที่ 2.6 แสดงถึงระดับของเลิร์นนิงอ็อบเจกต์นั้นเกิดจากการรวมสื่อข้อมูลย่อยหลายรูปแบบประกอบกันเป็นวัตถุสารสนเทศเพื่อแสดงแนวคิดของวัตถุประสงค์ย่อยบางเรื่อง จากนั้นจึงนำมาจัดเรียงและเชื่อมโยงลำดับการนำเสนอเกิดขึ้นเป็นสื่อการเรียนรู้เรื่องใดเรื่องหนึ่ง หรือเรียกว่าวัตถุประยุกต์ ทั้งนี้วัตถุประยุกต์แต่ละชิ้นก็จะถูกรวบรวมและจัดกลุ่มเนื้อหาที่สอดคล้องกันตามวัตถุประสงค์หลักเพื่อสร้างเป็นบทเรียน และเชื่อมโยงกันเป็นรายวิชาที่สมบูรณ์ในท้ายที่สุด

จากแนวคิดนิยามข้างต้น งานวิจัยนี้จึงใช้แนวคิดโครงสร้างองค์ประกอบของเลิร์นนิงอ็อบเจกต์ตามแบบของมาลิสัน (Mallison, 2001) มาประยุกต์ ซึ่งจำแนกองค์ประกอบออกเป็น 3 ระดับ ได้แก่ วัตถุสารสนเทศ (Information Object) วัตถุประยุกต์ (Application Object) และหน่วยบทเรียน (Lesson or Aggregates Assemblies) โดยเลิร์นนิงอ็อบเจกต์แต่ละรายการจะมีสื่อข้อมูลย่อย (Media Data Element) เป็นองค์ประกอบที่เล็กที่สุดในการสร้างเลิร์นนิงอ็อบเจกต์แต่ละรายการ ดังนั้นเพื่อให้สามารถจำแนกองค์ประกอบของเลิร์นนิงอ็อบเจกต์ลงไปถึงระดับที่เล็กที่สุด ที่สามารถนำกลับมาใช้ใหม่ได้ ในงานวิจัยนี้ได้ออกแบบโครงสร้างการจัดเก็บเลิร์นนิงอ็อบเจกต์ที่สามารถจำแนกเลิร์นนิงอ็อบเจกต์ตามระดับข้างต้น โดยรองรับการอธิบายส่วนประกอบย่อยของเลิร์นนิงอ็อบเจกต์แต่ละรายการที่ประกอบกันขึ้นมาเป็นเลิร์นนิงอ็อบเจกต์ในระดับที่สูงขึ้นได้ แต่ยังไม่รองรับการอธิบายถึงรูปแบบความสัมพันธ์และลำดับการประกอบเข้ากันระหว่างเลิร์นนิงอ็อบเจกต์ย่อย ๆ ได้ เนื่องจากข้อจำกัดในการสกัดองค์ประกอบย่อยของเนื้อหาเลิร์นนิงอ็อบเจกต์ในแต่ละประเภทที่มีความแตกต่างกัน รวมถึงต้องให้ผู้ที่เพิ่มข้อมูลเป็นผู้พิจารณารายละเอียดขององค์ประกอบแต่ละส่วนด้วยตนเอง จึงไม่สามารถให้ระบบทำงานแบบอัตโนมัติได้

2.3 การจัดเก็บเลิร์นนิงอ็อบเจกต์

จากข้อจำกัดของแหล่งข้อมูลบนอินเทอร์เน็ตที่เป็นแหล่งข้อมูลหลักในการค้นคืนและใช้งานเลิร์นนิงอ็อบเจกต์ของครูผู้สอน ได้แก่ การปิดกั้นการเข้าถึง รูปแบบการจัดเก็บเลิร์นนิงอ็อบเจกต์ (การกำหนดโครงสร้างข้อมูล หรือเมทาดาทา การให้คำค้น) ที่แตกต่างกันตามการออกแบบของผู้พัฒนาในแต่ละระบบ ซึ่งทำให้การค้นคืนเลิร์นนิงอ็อบเจกต์ให้สอดคล้องกับวัตถุประสงค์การนำไปใช้แต่ละระดับทำได้ยาก และใช้เวลามาก เพื่อแก้ปัญหาดังกล่าว งานวิจัยนี้จึงมุ่งเน้นพัฒนาแหล่งข้อมูลที่สามารถจัดเก็บและค้นค้นเลิร์นนิงอ็อบเจกต์ ที่รวบรวมมาจากแหล่งข้อมูลหลายแหล่งที่มีโครงสร้างเมทาดาทาที่แตกต่างกันมาจัดเก็บไว้ในโครงสร้างเดียวกันได้ และใช้แนวคิดของโอเพนดาทา (Open Data) ในการพัฒนาโครงสร้างการจัดเก็บข้อมูลของระบบแบบเปิดเสรี เพื่อสามารถเชื่อมโยงเขตข้อมูลไปยังแหล่งข้อมูลอื่น ๆ ที่มีมาตรฐานแบบเปิดเช่นเดียวกันได้ อีกทั้งยังช่วยเพิ่มความสามารถในการค้นคืน ให้ผู้ใช้สามารถค้นคืนจากแหล่งข้อมูลหลักผ่านไปยังแหล่งข้อมูลอื่น ๆ ในคราวเดียวกันได้

ในหัวข้อนี้เป็นการประมวลแนวคิดเกี่ยวกับการจัดการเลิร์นนิงอ็อบเจกต์ เพื่อนำไปใช้ในการกำหนดเค้ารางเมทาดาทา (Metadata Schema) ที่ที่มีความเหมาะสมสำหรับการอธิบายรายละเอียดของเลิร์นนิงอ็อบเจกต์ ให้สอดคล้องกับคุณลักษณะของเลิร์นนิงอ็อบเจกต์และพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของผู้ใช้ โดยอ้างอิงตามมาตรฐานเมทาดาทาที่เป็นมาตรฐานสากล และการออกแบบโครงสร้างของคลังคำเพื่อใช้ในการค้นคืนเลิร์นนิงอ็อบเจกต์

2.3.1 เมทาดาทา

คำว่า เมทาดาทา (Metadata) เป็นคำศัพท์ที่ถูกบัญญัติขึ้นมาเมื่อราวปลายศตวรรษที่ 1980 โดยคำว่า “Meta” มาจากภาษากรีกในความหมายว่า “about” หรือ “เกี่ยวกับ” ดังนั้น เมทาดาทาจึงมักหมายถึง ข้อมูลเกี่ยวกับข้อมูลอีกชุดหนึ่ง (Data About Data) หรือสารสนเทศเกี่ยวกับสารสนเทศอีกชุดหนึ่ง (Information about Information) เพื่อใช้บรรยายลักษณะ บอกตำแหน่ง หรือให้ข้อมูลอื่น ๆ เกี่ยวกับเอกสารหนึ่ง เพื่อให้การค้นคืนสารสนเทศ นำสารสนเทศไปใช้ และจัดการสารสนเทศทำได้ง่ายขึ้น (นิศาชล จำนงศรี, 2554: 77)

NISO (www, 2004) ได้อธิบายความหมายของเมทาดาทาว่าเป็นโครงสร้างข้อมูลที่ใช้อธิบาย ชี้ตำแหน่ง หรือใช้เพื่อการค้นคืน นำไปใช้ หรือจัดการทรัพยากรสารสนเทศ ซึ่งอาจจะเรียกว่าข้อมูลของข้อมูล (Data About Data) ทั้งนี้ยังชี้แจงถึงการให้คำจำกัดความของเมทาดาทาตามบริบทของกลุ่มนักวิชาชีพที่แตกต่างกัน เช่น กลุ่มนักเขียนโปรแกรมคอมพิวเตอร์หมายถึงสิ่งที่ใช้เพื่อให้ระบบคอมพิวเตอร์สามารถเข้าใจและแปลความหมายได้ในขณะที่ผู้ใช้ทั่วไปจะหมายถึงเขตข้อมูลที่อธิบายทรัพยากรสารสนเทศที่เป็นอิเล็กทรอนิกส์เท่านั้น กลุ่มบรรณารักษ์จะหมายถึงข้อมูลที่ใช้ในการลงรายการหรือเป็นตัวแทนทรัพยากรสารสนเทศโดยมีรูปแบบที่เป็นมาตรฐาน

อุทิตย์ พิมพา (2545: 28) ให้ความหมายในลักษณะเดียวกัน กล่าวคือ เมทาดาทาคือข้อมูลที่บอกรายละเอียดของข้อมูล ซึ่งอาจเป็นสิ่งของ กิจกรรม คน หรือหน่วยงาน และเมื่อมีการแปลงข้อมูลเหล่านี้ให้เป็นรูปดิจิทัลแล้ว ทุกสิ่งทุกอย่างจะมีสภาพเหมือนกันไม่สามารถแยกแยะรายละเอียดได้ จึงต้องมีข้อมูลที่เป็นรายละเอียดหรือตัวแทนของทรัพยากรดิจิทัลเหล่านั้น

ประดิษฐา ศิริพันธ์ (2546) ได้ให้ความหมายของเมทาดาทาคือ ข้อมูลที่มีการกำหนดโครงสร้างสำหรับอธิบายรายละเอียดของข้อมูลอิเล็กทรอนิกส์ที่บันทึกในฐานข้อมูล โดยใช้หลักเกณฑ์การทำข้อมูลเช่นเดียวกับการทำรายการบรรณานุกรมของหนังสือ เช่น อธิบายรายละเอียดข้อมูลชื่อเรื่อง ปีพิมพ์ ประเภทสิ่งพิมพ์ รูปแบบทางกายภาพ และข้อมูลเกี่ยวกับผู้แต่ง เป็นต้น

เวเบล (Weibel: 1997) ให้ความหมายของเมทาดาทา คือเขตข้อมูลที่บอกรายละเอียดของข้อมูล ซึ่งเป็นคำที่ใช้ทั่วไปในชุมชนอินเทอร์เน็ต สำหรับชุมชนห้องสมุดรู้จักกันในความหมายว่าเป็นการทำรายการหรือการลงรายการทรัพยากรสารสนเทศ หรือช่วยบอกให้ผู้ปฏิบัติงานเข้าใจว่า เป็นข้อมูลที่บอกรายละเอียดของหนังสือบนชั้น

ธอร์เนรี่ (Thornely: 1998) ให้คำจำกัดความของเมทาดาทาว่า เป็นข้อมูลที่ช่วยในการจัดระเบียบ พรรณนา แสดงลักษณะ ชี้ตำแหน่ง และค้นคืนทรัพยากรอิเล็กทรอนิกส์ หรือข้อมูลที่อยู่ในอินเทอร์เน็ตได้อย่างเหมาะสม

จากการนิยามความหมายของนักวิจัยและหน่วยงานที่เกี่ยวข้องข้างต้น สามารถสรุปคำจำกัดความของเมทาดาทาว่าเป็นชุดข้อมูลที่จัดทำขึ้นอย่างมีโครงสร้างเพื่อใช้ในการอธิบายทรัพยากรสารสนเทศที่อยู่ในรูปแบบอิเล็กทรอนิกส์ หรือสื่อดิจิทัล โดยมีหน้าที่เพื่อใช้เป็นตัวแทนสารสนเทศของทรัพยากรในแหล่งจัดเก็บเลิร์นนิงอ็อบเจกต์ และมีหน้าที่เพื่อใช้ในการกำหนดรายละเอียด ค้นคืน และแปลความหมายเพื่อสื่อความหมายให้ผู้ใช้เข้าใจรายละเอียดของทรัพยากรสารสนเทศ

2.3.1.1 ประเภทของเมทาดาทา

การสร้างเมทาดาทาจะแตกต่างกันไปขึ้นอยู่กับวัตถุประสงค์ของแต่ละหน่วยงาน โดยอาจจำแนกเมทาดาทาได้เป็น 3 ประเภท คือ (นิศาชล จำนงศรี, 2546: 82, Caplan, 2003)

1) เมทาดาทาเชิงบรรยาย (Descriptive Metadata) ได้แก่ เมทาดาทาที่ทำหน้าที่บรรยายรายละเอียดทางบรรณานุกรม และเนื้อหาของเอกสารเว็บ หรือเอกสารดิจิทัล เช่น ชื่อเรื่อง ชื่อผู้แต่ง ปีพิมพ์ หัวเรื่อง หรือคำสำคัญ เป็นต้น เช่นเดียวกับการบรรยายรายละเอียดทางบรรณานุกรมของวัสดุสารสนเทศในห้องสมุด

2) เมทาดาทาเชิงโครงสร้าง (Structural Metadata) ได้แก่ การอธิบายลักษณะโครงสร้าง หรือการจัดลำดับความสัมพันธ์ขององค์ประกอบต่าง ๆ ของเอกสารเว็บ หรือเอกสารดิจิทัล เช่น ลักษณะการเรียงลำดับหน้า จำนวนหน้า จำนวนบท/ตอน และการจัดเรียงลำดับ การจัดความสัมพันธ์ของหน้าต่าง ๆ ภายในเว็บไซต์ หรือการจัดโครงสร้างของเว็บไซต์ และการอธิบายประเภทของเอกสาร เช่น บทความ วารสาร รายงานการวิจัย สารานุกรม เป็นต้น เทียบได้กับการอธิบายลักษณะรูปร่างของทรัพยากรสารสนเทศในรายการสารสนเทศของห้องสมุด (Library Catalog) เนื่องจากโครงสร้างของเอกสารจะเป็นตัวบ่งบอกถึงลักษณะ หรือประเภทของเอกสาร

3) เมทาดาทาเพื่อการบริหาร (Administrative Metadata) ได้แก่ข้อมูลที่จำเป็นต่อการจัดการและการบำรุงรักษาข้อมูลดิจิทัลตลอดอายุการใช้งานของเอกสาร ประกอบด้วยข้อมูลหลัก 2 ส่วน คือ ส่วนข้อมูลเกี่ยวกับการสร้างเอกสาร เช่น การกำหนดค่าความละเอียด (Resolution) วันที่ทำการแปลงข้อมูล เป็นต้น เพื่อให้ผู้ใช้ทราบว่าจะต้องใช้ซอฟต์แวร์หรือฮาร์ดแวร์ใดบ้าง หากต้องการใช้เอกสาร และส่วนข้อมูลเกี่ยวกับการเข้าใช้เอกสาร ได้แก่ ข้อมูลด้านกฎหมายและการเงิน เช่น การแจ้งให้ทราบถึงผู้ที่มีสิทธิเข้าใช้ และค่าใช้จ่ายในการเข้าใช้ข้อมูล เป็นต้น โดยสามารถแบ่งเมทาดาทาประเภทนี้ออกเป็น 2 ประเภทย่อย ได้แก่

3.1) เมทาดาทาเชิงเทคนิค (Technical Metadata) เป็นเมทาดาทาเพื่อการสงวนรักษาและให้ข้อมูลเกี่ยวกับการสร้างทรัพยากรสารสนเทศ เช่น การกำหนดค่าความละเอียด (Resolution) ของข้อมูลดิจิทัล รูปแบบไฟล์ (File Format) วิธีการบีบอัดข้อมูล (Compression) วันที่ทำการแปลงข้อมูล เป็นต้น เพื่อให้ผู้ใช้ทราบว่าจะต้องใช้ซอฟต์แวร์หรือ

ฮาร์ดแวร์ใดบ้าง หรือวิธีกระบวนการเตรียมข้อมูลให้พร้อมใช้งานด้วยโปรแกรมหรือระบบสารสนเทศอื่น ๆ ที่เกี่ยวข้อง หากต้องการใช้ทรัพยากรสารสนเทศ

3.2) เมทาดาทาด้านสิทธิความเป็นเจ้าของ (Right Metadata) เป็นข้อมูลเกี่ยวกับการเข้าใช้ทรัพยากรสารสนเทศในเชิงลิขสิทธิ์หรือทรัพย์สินทางปัญญา เช่น การแจ้งให้ทราบถึงความเป็นเจ้าของ สิทธิในการเข้าถึงทรัพยากรสารสนเทศระดับต่าง ๆ สิทธิในการนำไปใช้ เป็นต้น

2.3.1.2 เมทาดาทาสำหรับบรรยายเลิร์นนิงอ็อบเจกต์

เค้าร่างเมทาดาทา (Metadata Schema) หรือชุดขององค์ประกอบเมทาดาทา (Set of Metadata Elements) สำหรับเลิร์นนิงอ็อบเจกต์ที่ได้รับความนิยมและถูกนำไปใช้ในงานในระบบการเรียนรู้อิเล็กทรอนิกส์ (e-Learning) และระบบคลังเลิร์นนิงอ็อบเจกต์ (Learning Object Repository) ได้แก่ ดับลินคอร์เมทาดาทา (Dublin Core Metadata) มาตรฐานลอม (Standard for Learning Object Metadata: LOM Standard) และมาตรฐานสคอร์ม (Sharable Content Object Reference Model: SCORM Standard) โดยมีรายละเอียดดังนี้

1) ดับลินคอร์ (Dublin Core: DC)

ดับลินคอร์เมทาดาทา (Dublin Core Metadata) นิยมใช้ตัวย่อ “DC” (DCMI, www, 2013) พัฒนาโดย Dublin Core Metadata Initiative (DCMI) ซึ่งประกอบด้วย บรรณารักษ์ นักเทคโนโลยีสารสนเทศและผู้ทำงานด้านโครงสร้างข้อมูลในเว็บ จากสหรัฐอเมริกา อังกฤษ ฝรั่งเศส ออสเตรเลีย และหลายประเทศในยุโรป เพื่อเป็นแบบแผนในการลงเมทาดาทาสำหรับเอกสารบนเว็บหรือทรัพยากรสารสนเทศที่ดิจิทัล ทั้งนี้ เพื่อช่วยในการค้นคืนสารสนเทศ ในปี ค.ศ. 1995 คณะทำงานดับลินคอร์ได้ประชุมกันครั้งแรกที่เมืองดับลิน รัฐโอไฮโอ และกำหนดชุดข้อมูลย่อย 15 หน่วย สำหรับใช้พรรณนาสารสนเทศดิจิทัลเพื่อให้เจ้าของผลงานจัดทำเมทาดาทาด้วยตนเอง และสามารถค้นคืนร่วมกันกับฐานข้อมูลต่างระบบ ปัจจุบันดับลินคอร์ได้รับการประกาศเป็นมาตรฐานสากล ISO 15836 และมาตรฐานของสหรัฐอเมริกา NISO Z39.85 ดับลินคอร์ได้รับการนำไปใช้อย่างกว้างขวางและเป็นพื้นฐานของการประยุกต์ใช้ ห้องสมุดต่าง ๆ (ประดิษฐา ศิริพันธ์: 2546)

จุดประสงค์ของดับลินคอร์ คือ การสร้างแบบแผนให้เจ้าของผลงานสามารถจัดทำเมทาดาทาได้ด้วยตนเอง โดยมีจุดเด่นทำให้ได้รับความนิยมคือ โครงสร้าง 15 องค์ประกอบข้อมูลย่อย การใช้งานไม่มีกฎเกณฑ์และรายละเอียดบังคับชัดเจน เพียงแต่เป็นการกำหนดให้เจ้าของผลงานเลือกใช้องค์ประกอบย่อยที่แสดงรายละเอียดของผลงานได้อย่างเหมาะสมและครบถ้วน จึงสามารถปรับใช้ได้กับข้อมูลหลายระบบ โดยองค์ประกอบข้อมูลย่อย 15 องค์ประกอบแบ่งได้เป็น 3 กลุ่มใหญ่ (Hillmann, 2001)

1.1) กลุ่มขององค์ประกอบที่เกี่ยวข้องกับเนื้อหาของทรัพยากรสารสนเทศ หมายถึง องค์ประกอบที่ใช้อธิบายเกี่ยวกับลักษณะของเนื้อหาทรัพยากรสารสนเทศนั้น ๆ ซึ่งมีองค์ประกอบที่เกี่ยวข้อง จำนวน 7 องค์ประกอบ ได้แก่ ชื่อเรื่อง (Title) หัวเรื่องหรือคำสำคัญ

(Subjects and Keywords) ลักษณะเนื้อหาสารสนเทศ (Description) ประเภท (Resource Type) ต้นฉบับ (Source) เรื่องที่เกี่ยวข้อง (Relation) และขอบเขต (Coverage)

1.2) กลุ่มขององค์ประกอบที่เกี่ยวข้องกับทรัพย์สินทางปัญญา หมายถึง องค์ประกอบที่ใช้อธิบายเกี่ยวกับสิทธิทางกฎหมาย เพื่อบ่งบอกถึงความเป็นเจ้าของทรัพยากรสารสนเทศนั้น ๆ ซึ่งมีองค์ประกอบที่เกี่ยวข้อง จำนวน 4 องค์ประกอบ ได้แก่ ผู้แต่งหรือเจ้าของผลงาน (Author or Creator) สำนักพิมพ์ (Publisher) ผู้ร่วมงาน (Contributor) และการจัดการด้านสิทธิต่าง ๆ (Rights)

1.3) กลุ่มขององค์ประกอบที่เกี่ยวข้องกับรูปแบบที่ปรากฏให้ใช้งาน หมายถึง องค์ประกอบที่ใช้อธิบายเกี่ยวกับลักษณะและการผลิต ซึ่งมีองค์ประกอบที่เกี่ยวข้อง จำนวน 4 องค์ประกอบ ได้แก่ ปี (Date) รูปแบบ (Format) รหัส (Resource Identifier) และภาษา (Language)

สำหรับการใช้งานดับลินคอร์เมทาดาทาในวงการการศึกษา โดยส่วนใหญ่มักจะถูกนำไปใช้ในระบบคลังเลิร์นนิงอ็อบเจกต์ (Learning Object Repository) เช่น ระบบค้นคืนทรัพยากรการเรียนรู้นานาชาติ (โครงการมหาวิทยาลัยไซเบอร์ไทย, www, 2556) และระบบคลังข้อมูลดิจิทัล (Digital Repository) ที่พัฒนาด้วยดีสเปส (DSpace) ซึ่งเป็นโอเพนซอร์ส (Open Source) ที่ผู้ใช้ส่วนหนึ่งนิยมนามาใช้สำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์ และ/หรือทรัพยากรสารสนเทศ เนื่องจากเป็นระบบที่มีความยืดหยุ่นและสามารถรองรับการจัดเก็บทรัพยากรสารสนเทศในรูปแบบที่หลากหลายได้ (DSpace, 2014) เป็นต้น

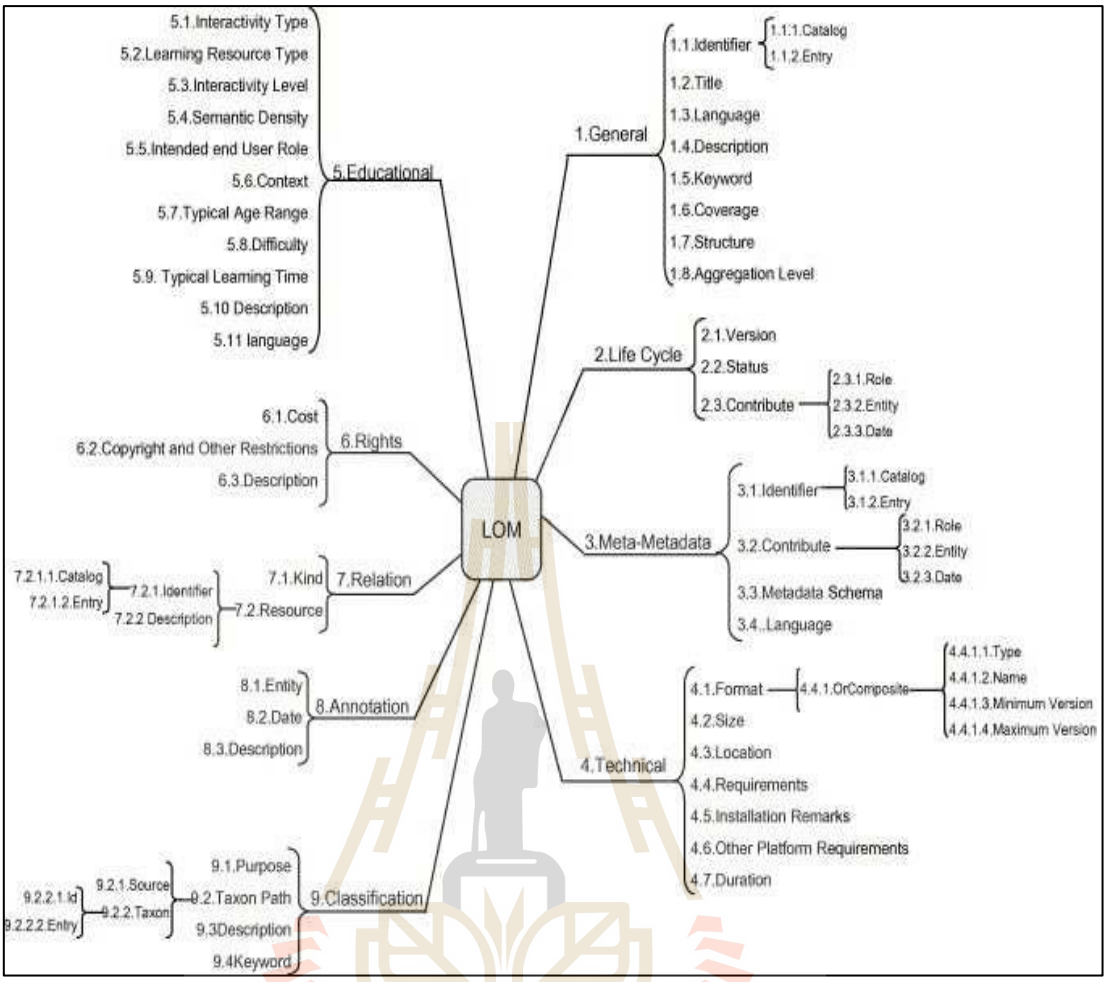
2) ลอม (Learning Object Metadata: LOM)

มาตรฐานลอมหรือ LOM Standard ย่อมาจากคำว่า Learning Object Metadata Standard เป็นมาตรฐานที่พัฒนาโดยไอทริปเปิลอี (IEEE) เพื่อเป็นมาตรฐานในการจัดเก็บเลิร์นนิงอ็อบเจกต์โดยเฉพาะ (IEEE, 2002) ที่จัดเก็บไว้ในคลังเลิร์นนิงอ็อบเจกต์ (Learning Object Repository) หรือระบบฐานข้อมูลในระบบจัดการสื่อการเรียนรู้ (Learning Object Management System) เพื่อให้ผู้ใช้สามารถเข้าถึงสื่อการเรียนรู้ดิจิทัลได้ง่ายและสะดวก ได้รับมาตรฐาน IEEE 1484.12-1-2002 ดังรูปที่ 2.7

โดยมีการแบ่งกลุ่มข้อมูลออกเป็น 9 คลาส (Class) ได้แก่ General, Life Cycle, Meta-Metadata, Technical, Educational, Rights, Relation, Annotation และ Classification สามารถอธิบายกลุ่มข้อมูลย่อยของเลิร์นนิงอ็อบเจกต์ตามมาฐานลอม ดังนี้

2.1) General เป็นคลาสที่ใช้อธิบายข้อมูลทั่วไปเกี่ยวกับแหล่งข้อมูลการเรียนรู้ทั้งหมด เช่น ชื่อเรื่อง ชื่อเรื่องรอง ภาษา คำดรรชนี เป็นต้น

2.2) Life Cycle เป็นคลาสที่ใช้อธิบายลักษณะของสถานะในอดีตและปัจจุบันของข้อมูลการเรียนรู้และผู้ที่เกี่ยวข้องในการพัฒนาส่วนของข้อมูลการเรียนรู้ หรือวงจรในการสร้างเนื้อหาและรายละเอียดของเลิร์นนิงอ็อบเจกต์ และรุ่น (Version) วัน เวลาในการพัฒนา เป็นต้น



รูปที่ 2.7 โครงสร้างและหน่วยข้อมูลย่อยของ LOM (IMS, 2004)

2.3) Meta-metadata เป็นคลาสที่ใช้อธิบายข้อมูลเกี่ยวกับคำอธิบายรายละเอียดของเลิร์นนิงอ็อบเจกต์อีกระดับหนึ่ง

2.4) Technical เป็นคลานที่ใช้อธิบายคุณลักษณะและความต้องการทางด้านเทคนิคของเลิร์นนิงอ็อบเจกต์

2.5) Educational เป็นคลาสที่ใช้อธิบายข้อมูลทางการศึกษาและลักษณะการเรียนการสอน เช่น ระดับการศึกษา ระดับความยากของเนื้อหา บทบาทของผู้เรียน ช่วงอายุที่เหมาะสมกับเนื้อหาภายในเลิร์นนิงอ็อบเจกต์ เป็นต้น

2.6) Rights เป็นคลาสที่ใช้อธิบายข้อมูลคุณสมบัติเกี่ยวกับทรัพย์สินทางปัญญาและเงื่อนไขในการนำไปใช้งาน

2.7) Relation เป็นคลาสที่ใช้อธิบายความสัมพันธ์ระหว่างแหล่งข้อมูลภายในเลิร์นนิงอ็อบเจกต์

2.8) Annotation เป็นคลาสสำหรับจัดเตรียมให้มีการแสดงความคิดเห็นเกี่ยวกับเนื้อหารายละเอียดภายในเลิร์นนิงอ็อบเจกต์

2.9) Classification เป็นคลาสอธิบายการกำหนดหมวดหมู่ให้กับทรัพยากรภายในระบบ

นอกจากกลุ่มข้อมูลหลักทั้ง 9 ประเภทแล้วมาตรฐานลอมยังมีการจำแนกรายละเอียดปลีกย่อยของกลุ่มข้อมูลทั้ง 9 เป็นองค์ประกอบข้อมูลย่อย (Element) เพิ่มเติมได้อีกทั้งนี้เพื่อการจัดเก็บรายละเอียดของสื่อการเรียนรู้ให้มีความละเอียดครบถ้วนและสามารถใช้เป็นข้อมูลตัวแทนของเลิร์นนิงอ็อบเจกต์ที่มีความสมบูรณ์ ผู้ใช้สามารถใช้ข้อมูลเหล่านี้ในการพิจารณาความสอดคล้องของเลิร์นนิงอ็อบเจกต์กับความต้องการของตนได้ง่ายยิ่งขึ้นมาตรฐานลอมจัดได้ว่าเป็นมาตรฐานเมทาดาทาในการจัดเก็บสื่อการเรียนรู้ที่มีความเหมาะสมในการอธิบายรายละเอียดของเลิร์นนิงอ็อบเจกต์ได้อย่างครบถ้วน

3) สคอร์ม (Sharable Content Object Reference Model: SCORM)

กิดานันท์ มลิทอง (2548) กล่าวถึง “SCORM” ว่าเป็นรูปแบบมาตรฐานในการแบ่งปันเนื้อหาที่ใช้ได้กับระบบบริหารจัดการเนื้อหา การเรียนรู้ หรือระบบจัดการเรียนรู้ (CMS) ที่ใช้กันทั่วไป และใช้ในการจัดระบบบัญชีรายชื่อและการทำเมทาดาทาอธิบายรายละเอียดของเลิร์นนิงอ็อบเจกต์ เพื่อให้ง่ายต่อการกำหนดตำแหน่ง และการประเมินความเหมาะสมของเนื้อหาภายในเลิร์นนิงอ็อบเจกต์ เช่น ชื่อเรื่อง ชื่อผู้แต่ง วันที่พิมพ์ จุดประสงค์ ลิขสิทธิ์ หัวเรื่อง หรือคำสำคัญ เป็นต้น และถูกรวบรวมอยู่บนระบบไฟล์ดิจิทัล โดยเก็บไว้ในคลังข้อมูล เพื่อให้เรียกใช้ได้ตลอดเวลา

ถนอมพร เลาหจรัสแสง (2550) กล่าวถึงมาตรฐานของเลิร์นนิงอ็อบเจกต์ว่ามาตรฐาน SCORM จะนำมาใช้เป็นองค์ประกอบสำคัญของทรัพยากรสารสนเทศที่เป็น e-Learning ทั้งด้านเนื้อหา (Content) และระบบบริหารจัดการเนื้อหา/สารสนเทศ (LMS/LCMS) โดยในด้านของเนื้อหา “Learning Object” ได้ถูกนำไปจัดกลุ่มหรือ Packaging ใหม่เพื่อได้รายละเอียดเกี่ยวกับระบบเพิ่มเติมในตัวเนื้อหา ตามข้อกำหนดของมาตรฐาน SCORM รายละเอียดต่าง ๆ จะถูกบันทึกในรูปของไฟล์ข้อมูล ซึ่งทำหน้าที่ อธิบายคุณสมบัติต่าง ๆ ของเลิร์นนิงอ็อบเจกต์นั้นร่วมกับมาตรฐานในระบบจัดการการเรียนรู้ โดยสคอร์มจะหยิบชุดเค้าร่างเมทาดาทาจากเมทาดาทาลอม (LOM) มาเป็นชุดเมทาดาทาหลัก และเพิ่มข้อกำหนดในการสร้างชุดข้อมูลและหีบห่อของชุดข้อมูล (Content Package) เพื่อรองรับการทำงานร่วมกับชุดคำสั่งของระบบการเรียนรู้อิเล็กทรอนิกส์ ตามข้อกำหนดของ IMS (EDUCAUSE Institutional Management System Project) เพื่อทำให้ข้อมูล

สามารถถ่ายโอนข้ามระบบจัดการการเรียนรู้ (LMS) รุ่นอื่น ๆ ได้โดยไม่ส่งผลกระทบต่อชุดข้อมูลภายในทรัพยากรสารสนเทศ ดังนั้นสคอร์มจึงเป็นมาตรฐานที่ใช้จัดเก็บเมทาดาทาของทรัพยากรสารสนเทศร่วมกับข้อกำหนดในการติดต่อระหว่างเนื้อหาและระบบจัดการการเรียนรู้ (LMS) เพื่อทำให้เลิร์นนิงอ็อบเจกต์สามารถใช้งานร่วมกันได้

จากมาตรฐานเมทาดาทาสำหรับเลิร์นนิงอ็อบเจกต์ที่กล่าวมาข้างต้น จะเห็นได้ว่าเมทาดาทาที่ดีจะต้องสามารถบรรยายลักษณะของทรัพยากรสารสนเทศได้อย่างครบถ้วนสอดคล้องกับวัตถุประสงค์การใช้งานและพฤติกรรมของผู้ใช้ ต้องไม่มีองค์ประกอบย่อยที่มากเกินไปจนยุ่งยากแก่การใช้งาน หรือน้อยเกินไปจนไม่ตอบวัตถุประสงค์การทำงาน องค์ประกอบย่อยจะต้องมีความชัดเจนใช้ง่าย งานวิจัยนี้ได้ทำการศึกษาเพื่อออกแบบเมทาดาทาเชิงบรรยาย (Descriptive Metadata) เพื่อใช้ในการจัดเก็บเลิร์นนิงอ็อบเจกต์ในแบบโอเพนดาทา และปรับใช้องค์ประกอบย่อยจากเค้าร่างเมทาดาทาที่เป็นที่ยอมรับในระดับสากล

2.3.2 คลังคำศัพท์สำหรับการค้นคืนสารสนเทศ

แลงแคสเตอร์ (Lancaster, 2003: 38) ให้ความหมายของคลังคำหรือคลังศัพท์ เป็นคำที่ใช้เรียกชุดคำศัพท์ที่จัดเก็บเป็นหมวดหมู่และมีความสัมพันธ์ของชุดคำ ระบบจัดเก็บและค้นคืนสารสนเทศจะใช้คลังคำเป็นตัวช่วยในการจัดทำดัชนีให้กับสารสนเทศที่จัดเก็บ และเป็นตัวช่วยของผู้ใช้ในการค้นคืนสารสนเทศ คลังคำจึงเป็นองค์ประกอบที่สำคัญของระบบจัดเก็บและค้นคืนสารสนเทศที่ช่วยให้ การค้นคืนด้วยดัชนีมีประสิทธิภาพมากยิ่งขึ้น

ไอชิซัน (Aichison, 2000: 1) ได้ให้ความหมายของคลังคำศัพท์ว่า เป็นคำศัพท์ที่ทำหน้าที่ควบคุมภาษาดัชนีและการจัดการตามระเบียบวิธี ซึ่งเรียงตามลำดับความสัมพันธ์ของคำ โดยใช้ในระบบจัดเก็บและค้นคืนสารสนเทศ การจัดเรียงจากบัตรรายการสู่อินเทอร์เน็ต

นฤมล ปราชญโยธิน (2556: 90) ได้ให้ความหมายของศัพท์สัมพันธ์ไว้ว่า เป็นคลังคำศัพท์ควบคุมที่ผู้สร้างได้กำหนด หรือแนะให้ใช้ศัพท์คำใดคำหนึ่งสำหรับคำหรือกลุ่มคำที่มีความหมายเหมือนกัน และประมวลศัพท์ที่อยู่ในกลุ่มหรือสกุลเดียวกันไว้ด้วยกัน อีกทั้งยังแสดงคำที่เกี่ยวข้องสัมพันธ์กันในลักษณะอื่น ๆ ไว้ด้วย โดยมีสัญลักษณ์แสดงความสัมพันธ์ของคำเหล่านั้นกำกับไว้เพื่อนำไปใช้เป็นคำแทนสาระหรือ คำดัชนีของเอกสารในขั้นตอนการจัดเก็บเอกสารเข้าระบบ และทำหน้าที่เป็นคำค้นในขั้นตอนการค้นคืนเอกสาร คำอื่น ๆ ที่ใช้เรียกคลังคำ อาทิ ศัพท์สัมพันธ์ อรรถาภิธานศัพท์ หรือธิซอรัส เป็นต้น

สามารถสรุปความหมายของคลังคำได้ว่า หมายถึง ชุดคำศัพท์ที่ถูกจัดเก็บไว้เป็นหมวดหมู่และมีการกำหนดลักษณะโครงสร้างความสัมพันธ์ของคำศัพท์ในชุด โดยชุดคำศัพท์เป็นคำศัพท์ควบคุม (Control Vocabulary) ที่มีการกำหนดจากผู้เชี่ยวชาญหรือแหล่งข้อมูลในสายวิชาชีพนั้น ๆ โดยชุดคำศัพท์เหล่านี้จะถูกนำไปใช้เป็นตัวแทนแทนสาระหรือคำดัชนีของเอกสารในขั้นตอนการจัดเก็บสารสนเทศ และทำหน้าที่เป็นคำค้นในขั้นตอนการค้นคืนสารสนเทศ

นฤมล ปราชญโยธิน, ทวีศักดิ์ กออนันตกูล และเปรมิน จินดาวิมลเลิศ. (2536: 53-54) อธิบายหน้าที่ทั้ง 2 ด้านของคลังคำไว้ ดังนี้

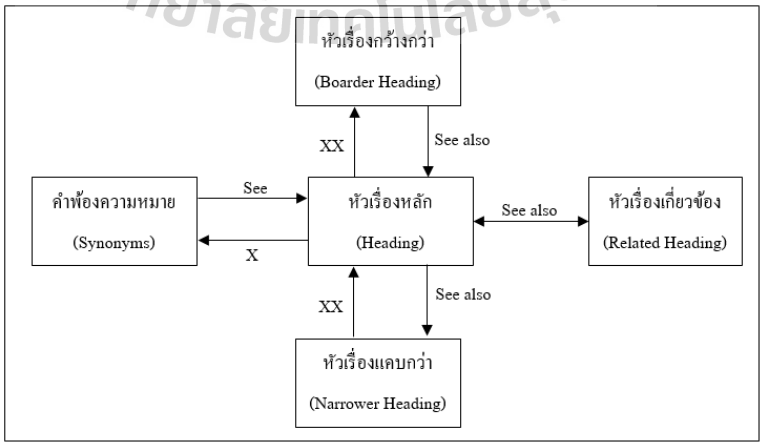
1) บทบาทในขั้นตอนการจัดเก็บ ในขั้นนี้คลังคำจะทำหน้าที่เป็นคู่มือของนักทำดัชนีในระบบภาษาควบคุม กล่าวคือ นักทำดัชนีจะวิเคราะห์เนื้อหาของเอกสาร แล้วแจกแจงผลการวิเคราะห์ เนื้อเรื่อง และแง่มุมของเอกสารนั้นออกมาเป็นภาษาดัชนี โดยอาศัยคลังคำศัพท์เป็นหลักในการเลือกใช้คำศัพท์ ให้มีความหมายพอเหมาะพอดีกับเนื้อหาที่แท้จริงของเอกสารและควบคุมความคงเส้นคงวาในการใช้คำให้ตรงกันทุกครั้งที่ต้องกำหนดคำแทนสาระให้กับเอกสารที่มีเนื้อหาหรือแง่มุมเดียวกัน ทั้งนี้เพื่อให้เกิดประสิทธิภาพและประสิทธิผลในขั้นตอนการค้นคืนต่อไป

2) บทบาทในขั้นตอนการค้นคืน ผู้ค้นคืนจะอาศัยคลังคำศัพท์เป็นคู่มือในการเสาะหาคำที่ตรงกับความหมายของเรื่องที่ตนจะค้นคืน กระบวนการเริ่มด้วยการวิเคราะห์และตีความคำถามแล้วแปลงเนื้อเรื่องหรือแง่มุมของคำถามนั้นออกมาเป็นภาษาดัชนี ซึ่งเป็นหลักการเดียวกับขั้นตอนการวิเคราะห์เอกสารเพื่อการจัดเก็บนั่นเอง

2.3.2.1 โครงสร้างความสัมพันธ์ในคลังคำ

การแสดงโครงสร้างความสัมพันธ์ของคำศัพท์ในลักษณะของศัพท์ควบคุม ที่นิยมนำมาใช้อย่างแพร่หลายในการจัดเก็บและค้นคืนสารสนเทศ ได้แก่ ศัพท์ควบคุมประเภทหัวเรื่อง (Subject Heading) และศัพท์ควบคุมประเภทธิซอรัส (Thesaurus) ซึ่งมีลักษณะโครงสร้างความสัมพันธ์ของคำศัพท์ในคลังคำ ดังนี้

1) โครงสร้างความสัมพันธ์ของศัพท์ควบคุมประเภทหัวเรื่อง ซึ่งโดยทั่วไปมีหลักการระบุความสัมพันธ์ของคำศัพท์ในระดับที่กว้างกว่าและแคบกว่าและเชื่อมโยงจากคำที่ไม่ใช้ไปยังคำที่ใช้แทนได้ (ดังรูปที่ 2.8) โดยแสดงด้วยสัญลักษณ์ ดังนี้



รูปที่ 2.8 โครงสร้างความสัมพันธ์ของศัพท์ควบคุมประเภทหัวเรื่อง (นฤมล ปราชญโยธิน, 2556)

See (ดูที่) ใช้ในกรณีโยงคำที่ไม่ใช้ไปสู่คำที่ใช้แทน

See also (ดูเพิ่มเติม) ใช้ในกรณีโยงไปสู่คำที่สัมพันธ์กัน แต่มีความหมายของคำที่แคบกว่าคำหลัก

X ใช้ระบุให้ทราบว่าคานั้นเป็นคำที่ไม่ใช้

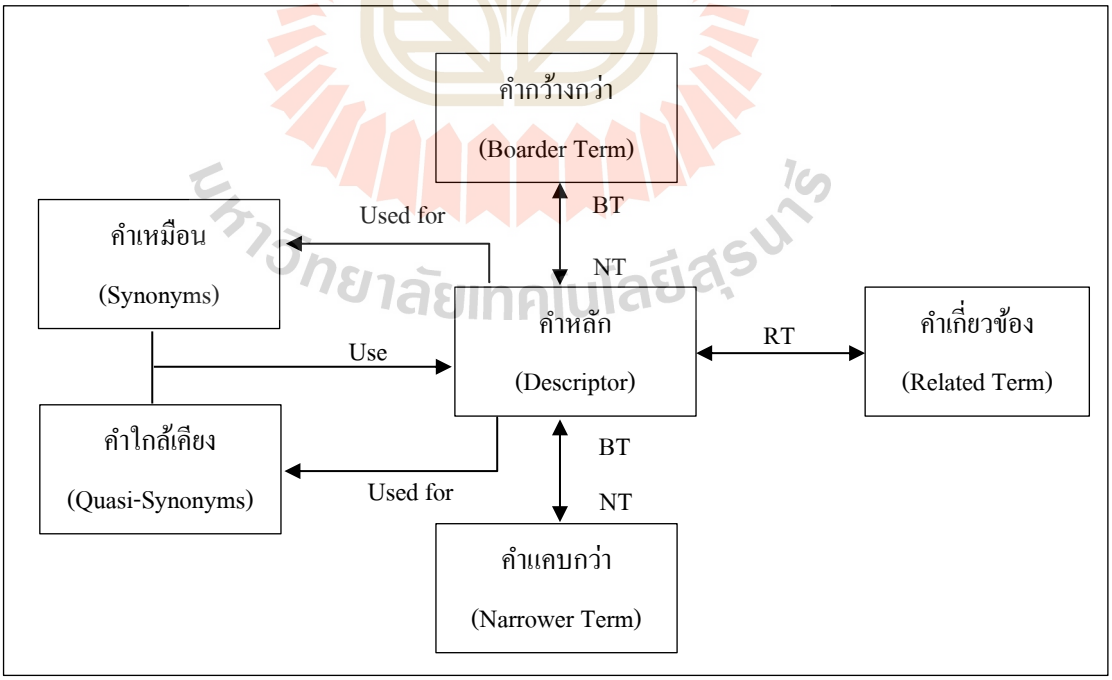
XX ใช้ในกรณีโยงไปสู่คำที่สัมพันธ์กัน แต่เป็นคำที่มีความหมายกว้างกว่าคำหลัก

2) โครงสร้างความสัมพันธ์ของธิซอรัสเป็นการรวบรวมคำศัพท์หรือบัญชีคำศัพท์เป็นชุด ๆ แต่ละชุดมีส่วนประกอบทั่ว ๆ ไป ตามรูปที่ 2.9 ดังนี้ (นฤมล ปราชญโยธิน ทวีศักดิ์ กออนันตกูล และเปรมิน จินดาวิมลเลิศ, 2536: 54)

2.1) คำหลักของชุด (Descriptor)

2.2) คำที่มีความหมายพ้องกับคำหลักหรือคำที่มีความหมายใกล้เคียงคำหลัก แต่ระบบไม่ใช้และไม่ได้กำหนดให้ใช้คำหลักแทน (Used for)

2.3) ข้อความอธิบายคำหลัก (Scope Note) เพื่อให้เกิดความกระจ่างเกี่ยวกับศัพท์คำนั้น แต่ไม่จำเป็นว่าศัพท์ทุกคำจะต้องมีคำอธิบาย



รูปที่ 2.9 โครงสร้างความสัมพันธ์ของคำศัพท์ควบคุมประเภทธิซอรัส

2.4) คำที่มีความสัมพันธ์กับคำหลัก ได้แก่ คำที่เป็นต้นสกุลของคำหลัก (Top Term) คำที่มีความหมายกว้างกว่าคำหลัก (Broader Term) คำที่มีความหมายแคบกว่าคำหลัก (Narrower Term) คำที่มีความหมายเกี่ยวข้องกับคำหลัก (Related Term) แต่ไม่ใช่ในลักษณะที่เป็นคำที่มีความหมายกว้างกว่าหรือแคบกว่า

SN (Scope note)	ใช้นำหน้าข้อความอธิบายคำศัพท์
U (Use)	ใช้นำหน้าคำศัพท์ที่กำหนดให้ใช้
UF (Used for)	ใช้นำหน้าคำศัพท์ที่ไม่ใช้
TT (Top term)	ใช้นำหน้าคำศัพท์ที่ต้นสกุล หรือคำรวม
BT (Broader term)	ใช้นำหน้าคำศัพท์ที่มีความหมายกว้างกว่าคำหลัก
NT (Narrower term)	ใช้นำหน้าคำศัพท์ที่มีความหมายแคบกว่าคำหลัก
RT (Related term)	ใช้นำหน้าคำศัพท์ที่มีความหมายเกี่ยวข้องกับคำหลัก

งานวิจัยนี้เลือกใช้วิธีการจัดทำโครงสร้างความสัมพันธ์ของคลังคำแบบธิซอรัส ในการจัดทำคลังคำศัพท์ทางคณิตศาสตร์ในระดับมัธยมศึกษาแบบสองภาษา คือ ภาษาไทยและภาษาอังกฤษ เนื่องจากสามารถอธิบายความสัมพันธ์ได้ครบถ้วน ทั้งความสัมพันธ์แบบกว้างกว่า แคบกว่า และเกี่ยวข้องกัน คำพ้องความหมาย และสอดคล้องกับโครงสร้างของสคอส (Simple Knowledge Organization System: SKOS) ซึ่งเป็นโครงสร้างในการพัฒนาคลังคำเพื่อสนับสนุนการค้นคืนเชิงความหมาย

2.3.2.2 กระบวนการสร้างคลังคำ

แนวทางการพัฒนาธิซอรัส มี 3 แนวทาง ดังต่อไปนี้ (นฤมล ปราชญโยธิน, ทวีศักดิ์ กออนันตกูล และเปรมิน จินดาวิมลเลิศ. 2536: 72)

1) การปรับใช้ธิซอรัสที่มีอยู่แล้ว คือ การปรับหรือขยายขอบเขตของธิซอรัสที่มีอยู่แล้วนั้นให้เหมาะแก่การนำไปใช้ของหน่วยงาน

2) การแปลธิซอรัสที่มีอยู่แล้ว คือ การนำธิซอรัสที่มีอยู่แล้ว มาทำการแปลเป็นภาษาที่เหมาะสมแก่การใช้งาน

3) การสร้างธิซอรัสใหม่ คือ การสร้างธิซอรัสขึ้นมาใหม่ เนื่องจากไม่มีธิซอรัสในสาขาวิชานั้นเลย

กระบวนการพัฒนาธิซอรัสใหม่ ประกอบด้วย 9 ขั้นตอน ดังนี้ (นฤมล ปราชญโยธิน ทวีศักดิ์ กออนันตกูล และเปรมิน จินดาวิมลเลิศ. 2536: 73-77)

1) การวางแผนดำเนินงาน ในด้านเนื้อหาและรูปแบบ รวมไปถึงการกำหนดขอบเขต ความครอบคลุมไปถึง ขอบข่ายของการใช้ มาตรฐานการควบคุมศัพท์ และวิธีการทาง

ไวยากรณ์ที่เหมาะสม เป็นต้น ส่วนด้านทรัพยากรในการดำเนินงาน เน้นการจัดสรรกำลังคนการแสวงหา และทาบทามผู้ทรงคุณวุฒิ การเตรียมวัสดุอุปกรณ์เพื่อการดำเนินงาน เป็นต้น

2) การสะสมคำศัพท์ คำศัพท์อาจได้มาจากแหล่งต่าง ๆ ดังต่อไปนี้

2.1) จากธิซอรัสที่มีเผยแพร่อยู่ในขณะนั้น โดยเลือกธิซอรัสที่มีขอบเขตเนื้อหาครอบคลุมวงเรื่องที่จะจัดสร้าง เพื่อบันทึกชุดคำศัพท์เฉพาะแขนงที่ต้องการหรือที่เกี่ยวข้องออกมา

2.2) จากสิ่งพิมพ์ทั่วไปในแขนงวิชานั้น เช่น เอกสารการวิจัย รายงานการประชุมทางวิชาการ หนังสือ วารสาร และอื่น ๆ

2.3) จากบรรณานุกรมเฉพาะวิชา โดยเฉพาะอย่างยิ่งบรรณานุกรมที่มีสาระสังเขปด้วย

2.4) จากศัพท์ดรรชนี ที่ปรากฏในเอกสารประเภทบริการสาระสังเขปและดรรชนี

2.5) จากหัวเรื่องต่าง ๆ

2.6) จากอภิธานศัพท์ และพจนานุกรมเฉพาะแขนง

2.7) จากแนวแนะของผู้เชี่ยวชาญเฉพาะสาขา

3) การประมวลคำศัพท์ คือ การนำคำศัพท์ที่ได้จากขั้นตอนการสะสมคำศัพท์เข้าไว้ด้วยกันอย่างเป็นระบบ วัตถุประสงค์ของขั้นตอนนี้ คือ

3.1) เพื่อขจัดคำซ้ำที่ได้มาจากแหล่งต่าง ๆ

3.2) เพื่อรวมคำพ้องความหมาย หรือคำใกล้เคียงเข้าด้วยกันผลลัพธ์ของขั้นตอนนี้ จะได้คำศัพท์ 2 กลุ่มคือกลุ่มที่ 1 เป็นคำศัพท์ที่ใช้ในการสร้างธิซอรัสกลุ่มที่ 2 เป็นคำศัพท์ที่จะไม่ใช้ ซึ่งต้องทำรายการเชื่อมโยง (U/UF) ไว้

4) จัดกลุ่มคำศัพท์และจัดประเภท ลำดับของความสัมพันธ์ของศัพท์ในแต่ละกลุ่ม

5) ปรับ/เพิ่มคำศัพท์ที่เกี่ยวข้องในลักษณะต่าง ๆ (TT BT NT RT U/UF)

6) การตรวจสอบและปรับฐานะของคำศัพท์ เพื่อบันทึกคำศัพท์ในลำดับความสัมพันธ์ถัดไป ที่ยังมิได้อยู่ในฐานะคำหลัก ลงในบัตรเพิ่ม เพื่อให้ศัพท์ทุกคำมีโอกาสเรียงอยู่ในตำแหน่งคำหลัก ทั้งนี้จะต้องบันทึกคำที่สัมพันธ์กับศัพท์นั้นลงไปด้วย

7) การตรวจตราและปรับแต่งรายการคำศัพท์ แต่ละชุดให้ถูกต้องสมบูรณ์และมีความสอดคล้องกันในระหว่างคำศัพท์ที่สัมพันธ์กันในลักษณะใดลักษณะหนึ่ง

8) การเรียงและจัดแสดงรายการคำศัพท์ รายการคำศัพท์ที่ควรให้มีความแตกต่างระหว่างคำศัพท์ที่เลือกใช้และไม่เลือกใช้ เช่น ในธิซอรัสฉบับพิมพ์ คำศัพท์ที่ใช้ จะพิมพ์ด้วยตัวหนา คำศัพท์ที่ไม่ใช้ จะพิมพ์ด้วยตัวบาง หรือพิมพ์อยู่ในรูปของตัวพิมพ์ใหญ่กับตัวพิมพ์เล็กตามลำดับ เป็นต้น ทั้งนี้เพื่อให้ผู้ใช้สังเกตเห็นได้อย่างชัดเจน เพื่อความสะดวกรวดเร็วในการค้นหาคำศัพท์ที่ต้องการ

9) การตรวจ/ เรียงลำดับอักษรของคำศัพท์ที่อยู่ภายใต้ BT NT RT และจัดแสดงรายการคำศัพท์

2.3.2.3 แหล่งคำศัพท์ในคลังคำ

แหล่งที่มาของคำศัพท์ในคลังคำจำแนกออกเป็น3 แหล่ง (นฤมล ปราชญโยธิน, 2556: 33-36, 47-48) คือ

1) แหล่งคำ เป็นแหล่งข้อมูลหรือสื่อทรัพยากรสารสนเทศในสาขาวิชาที่เกี่ยวข้อง เป็นแหล่งรวมของคำต่าง ๆ ที่สามารถนำมาใช้ในการสร้างคลังคำ ปกติแล้วแหล่งข้อมูลที่เป็นแหล่งคำเหล่านี้จะมีมาตรฐานในการจัดเก็บคำที่เป็นคำศัพท์ควบคุม หรือมีแบบแผนในการจัดเก็บอย่างเป็นทางการ โดยสามารถนำไปพัฒนาเป็นคลังคำได้ค่อนข้างสมบูรณ์ แหล่งที่มาของคำที่เป็นแหล่งคำ ได้แก่ พจนานุกรม (Dictionary) แหล่งภาษาดัชนี (Indexing) แผนการจัดหมวดหมู่เอกสาร (Catalog) คำค้นหรือคำถาม (Query) และภาคส่วนที่ว่าด้วยคำศัพท์ในวารสารหรือหนังสือพิมพ์ เป็นต้น

2) แหล่งความ เป็นแหล่งข้อมูลที่จะต้องอาศัยการอ่านและสกัดคำออกมาจากเนื้อความนั้น ๆ ของแหล่งความ ทั้งนี้ขึ้นอยู่กับผู้สร้างคลังคำ ว่ามีความต้องการที่จะสกัดคำออกมาจากเนื้อความของทรัพยากรสารสนเทศประเภทใด หรือจากแหล่งความรู้ในเรื่องใด สำหรับแหล่งความที่จัดได้ว่ามีประโยชน์มากสำหรับผู้สร้างคลังคำคือ ชื่อเรื่องของทรัพยากรสารสนเทศต่าง ๆ เพราะโดยส่วนใหญ่มักจะมีการตั้งชื่อเรื่องอย่างตรงไปตรงมา นอกจากนี้ ยังมีแหล่งความต่าง ๆ ที่สามารถสกัดคำออกมาจากเนื้อความที่เกี่ยวข้องหรือสนใจ ได้แก่ สารานุกรม สาระสังเขป รวมถึงเอกสารต่าง ๆ เช่น หนังสือ ตำรา วารสาร หรือนิตยสาร เป็นต้น

3) แหล่งบุคคล จะหมายรวมถึง ผู้ทรงคุณวุฒิ ผู้เชี่ยวชาญและผู้ใช้ ซึ่งความรู้และความคิดเห็นจากบุคคลต่าง ๆ ที่มีส่วนเกี่ยวข้องกับคลังคำ นอกจากจะช่วยในแง่ของการเป็นแหล่งคำและโครงสร้างแนวคิดเกี่ยวกับเรื่องที่ต้องการสร้างคลังคำแล้ว การช่วยตรวจสอบที่มาของคำ การวิจารณ์ การวิพากษ์และ/หรือการตั้งคำถามต่าง ๆ เป็นสิ่งที่ก่อให้เกิดการสร้างสรรค์ของกระบวนการสร้างคลังคำด้วย

จากแนวคิดดังกล่าวข้างต้น ในงานวิจัยนี้ได้สร้างชุดคำศัพท์ขึ้นจากแหล่งที่มาทั้ง 3 แหล่งข้างต้น โดยมีรายละเอียดดังนี้

1) แหล่งคำ โดยการเก็บรวบรวมคำศัพท์ในรายวิชาคณิตศาสตร์ จากแหล่งที่มีการจัดหมวดหมู่ชุดคำศัพท์เอาไว้แล้ว ได้แก่

1.1) แบบแผนการจำแนกหมวดหมู่ความรู้ระบบทศนิยมของดิวอี้ (Dewey Decimal Classification System) ซึ่งเป็นการจำแนกความรู้ออกเป็น 10 หมวดใหญ่ และภายในแต่ละหมวดจะแบ่งออกได้อีก 10 หมวดย่อย และในแต่ละหมวดย่อยจะจำแนกออกเป็น 10 หมู่ย่อย ถ้ามีเนื้อหาที่แบ่งย่อยลงไปอีกสามารถใส่จุดทศนิยมได้ไม่มีสิ้นสุด ทั้งนี้ในกลุ่มสาระการเรียนรู้คณิตศาสตร์จะอยู่ในหมวดหมู่หลัก 500 วิทยาศาสตร์ และในหมวดหมู่ย่อย 510 คณิตศาสตร์ ซึ่งหัวเรื่องในระบบดิวอี้นั้น มีความเหมาะสมเพียงพอต่อการจัดทำโครงสร้างคลังคำ เนื่องจากคำศัพท์ทางคณิตศาสตร์ในระดับมัธยมไม่มีความซับซ้อนมากนัก ดังนั้นในงานวิจัยนี้จึงใช้หมวดหมู่ DDC 510 ในการวิเคราะห์โครงสร้างการจัดหมวดหมู่และคำศัพท์ที่เหมาะสมถูกต้องในหัวเรื่องคณิตศาสตร์ระดับมัธยมศึกษา

1.2) เวิร์ดเน็ต (WordNet) เป็นคลังคำ หรือแหล่งคำศัพท์ในหลายสาขาวิชา ซึ่งรวมทั้งคำศัพท์ทางคณิตศาสตร์ เวิร์ดเน็ตมีโครงสร้างความสัมพันธ์ของคำคล้ายคลึงกับแบบธิซอรัส ทำให้การเชื่อมโยงคำศัพท์ที่เกี่ยวข้องมีรายละเอียดที่ครอบคลุมและหลากหลายมิติ งานวิจัยนี้ได้นำชุดคำศัพท์ทางคณิตศาสตร์จากเวิร์ดเน็ตมาใช้ในการกำหนดโครงสร้างของคลังคำศัพท์ทางคณิตศาสตร์ร่วมกับมาตรฐานจัดหมวดหมู่ระบบทศนิยมของดิวอี้ เพื่อช่วยให้หัวเรื่องและโครงสร้างมีความสมบูรณ์มากยิ่งขึ้น

1.3) พจนานุกรมศัพท์ทางคณิตศาสตร์ เป็นแหล่งคำศัพท์ที่มีความถูกต้องและครบถ้วนมากที่สุดในการสร้างคลังคำ งานวิจัยนี้ได้กำหนดคลังคำศัพท์ที่สามารถรองรับทั้งภาษาไทยและภาษาอังกฤษ ดังนั้น จึงต้องมีการใช้คำศัพท์ทางคณิตศาสตร์เพื่อช่วยให้การกำหนดหัวเรื่อง และคำศัพท์ทั้งสองภาษามีความถูกต้องและสัมพันธ์กันตามโครงสร้างคลังคำ

2) แหล่งความ โดยทำการเก็บรวบรวมคำศัพท์และสกัดคำศัพท์ต่าง ๆ จากแหล่งความที่มีเนื้อหาของกลุ่มสาระการเรียนรู้คณิตศาสตร์ สอดคล้องกับหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 โดยสกัดจากชื่อเลิร์นนิงอ็อบเจกต์ที่เป็นสื่อการสอนในเว็บไซต์ของกระทรวงศึกษาธิการ สำนักงานส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี และเว็บไซต์การศึกษาทางไกล (DLIT) ซึ่งสื่อการสอนเหล่านี้ได้จัดทำและจัดกลุ่มเนื้อหาตามโครงสร้างหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานอยู่แล้ว ทำให้สามารถแบ่งกลุ่มเนื้อหาของสื่อการสอนที่จัดเก็บได้เหมาะสมยิ่งขึ้น นอกจากนี้ ยังวิเคราะห์รายละเอียดจากหนังสือเรียนตามหลักสูตร

แกนกลางฯ ที่กระทรวงศึกษาธิการเป็นผู้ผลิตโดยตรง เพื่อสกัดคำศัพท์ในเนื้อหาบทเรียน และคำดัชนีท้ายเล่มในแต่ละบท เพื่อให้มีคลังคำศัพท์ที่ครอบคลุมเนื้อหา และเป็นคำศัพท์ที่สอดคล้องกับการใช้งานจริงในหลักสูตรภาษาไทย และภาษาอังกฤษ

การพัฒนาธิซอรัสจะพัฒนาตามมาตรฐานสคอส (SKOS) ซึ่งสามารถอธิบายความหมายของคำศัพท์ตามโครงสร้างของธิซอรัสได้ และจะใช้โปรแกรม Protégé ในการจัดเก็บรายละเอียดและความสัมพันธ์ของคำต่าง ๆ ตามโครงสร้างของสคอสทั้งนี้เพื่อสามารถนาไปใช้ในการพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

2.3.3 ลิงค์โอเพนดาทา (Linked Open Data)

ปัจจุบันหน่วยงานและสถาบันต่าง ๆ เริ่มมุ่งเน้นให้ระบบสารสนเทศบนอินเทอร์เน็ตพัฒนาในรูปแบบระบบเปิดมากขึ้น ซึ่งแสดงถึงแนวคิดการส่งเสริมการแลกเปลี่ยนข้อมูล และการใช้ทรัพยากรสารสนเทศร่วมกันอย่างเสรี รูปแบบการจัดเก็บเลิร์นนิงอ็อบเจกต์ในงานวิจัยนี้จึงใช้แนวคิดการจัดเก็บบนมาตรฐานแบบเปิดเสรีได้แก่ ลิงค์โอเพนดาทา (Linked Open Data) ซึ่งเป็นมาตรฐานในการกำหนดโครงสร้างการจัดเก็บข้อมูลที่สามารถเข้าถึง แลกเปลี่ยน เก็บเกี่ยวข้อมูลระหว่างกันได้อย่างเสรี โดยมีรายละเอียดดังต่อไปนี้

2.3.3.1 ความหมายของลิงค์โอเพนดาทา

ลิงค์โอเพนดาทาเป็นการนำแนวคิดการเชื่อมโยงข้อมูล หรือลิงค์ดาทา (Linked Data) ผสานกับแนวคิด การเปิดเผยข้อมูลแบบเสรี (Open Data) เพื่อให้ข้อมูลที่ไม่มีข้อผูกมัดในการใช้งาน สามารถเชื่อมโยงเข้าเป็นเครือข่ายและเข้าถึงได้สะดวก (Berners-Lee, 2006)

โดยแนวคิดของลิงค์ดาทา จะมุ่งเน้นการเผยแพร่ข้อมูลที่มีโครงสร้างเพื่อให้ข้อมูลนั้นสามารถเชื่อมโยงข้อมูลที่เกี่ยวข้องกันต่อ ๆ กันได้แบบอัตโนมัติ โดยอาศัยเทคโนโลยีเว็บได้แก่ โปรโตคอลเอชทีทีพี (HTTP) โครงสร้างการอธิบายข้อมูลแบบอาร์ดีเอฟ (Resource Description Framework) และ URI (Uniform Resource Identifier) ซึ่งทำให้ข้อมูลที่มาจากแหล่งข้อมูลต่างกันสามารถเชื่อมต่อกันได้

ขณะที่แนวคิด การเปิดเผยข้อมูล (Open Data) มีแนวคิดว่า ข้อมูลบางอย่างที่สามารถเปิดเผยเผยได้ ควรสามารถเปิดให้ทุกคนเข้าถึง ใช้งาน และนำกลับมาใช้ใหม่ได้อย่างเสรี โดยไม่มีข้อจำกัดด้านลิขสิทธิ์ สิทธิบัตร หรือเงื่อนไขทางเทคนิคอื่น ๆ

1) มาตรฐานการจัดเก็บข้อมูลในลิงค์โอเพนดาทา

หลักการในการสร้างเอกสารด้วยลิงค์โอเพนดาทามีรายละเอียดดังต่อไปนี้

1.1) การกำหนดชื่อของรายการเอกสารด้วยยูอาร์ไอ (Uniform Resource Identifier: URI) เปรียบเหมือนกับชื่อของเอกสาร โดยกำหนดเป็นโครงสร้างของการจัดเก็บเนื้อหาภายในคลังข้อมูลหรือภายในเว็บไซต์ เช่น การเก็บข้อมูลของบุคคล โดยแบ่งเป็นข้อมูลพนักงานและลูกค้า สามารถออกแบบยูอาร์ไอได้ดังนี้ yoursite/people/employee/department1/NameofPerson สำหรับยูอาร์ไอของลูกค้าได้แก่ yoursite/people/customer/CustomerName ซึ่งระดับการจัดเก็บผู้ใช้สามารถออกแบบได้เองตามความเหมาะสมในการจัดหมวดหมู่ของข้อมูลภายในฐานข้อมูลหรือชุดข้อมูลในเว็บไซต์ เพียงแต่ต้องพิจารณาถึงการขยายขอบเขตการจัดเก็บ

1.2) ใช้มาตรฐานการสื่อสารผ่านอินเทอร์เน็ตด้วยเอชทีทีพี (HTTP Protocol) เมื่อจัดเก็บชื่อเอกสารในโครงสร้างแบบยูอาร์ไอแล้ว เพื่อให้เอกสารสามารถเข้าถึงผ่านระบบอินเทอร์เน็ตจึงนิยมใช้โพรโทคอลเอชทีทีพี (HTTP Protocol) ร่วมด้วย เช่น http://yoursite/../yourdocument.html เป็นต้น รวมถึงการตั้งชื่อที่สามารถจัดการดูแลได้ง่ายแต่ไม่มีโอกาสที่จะซ้ำกัน เพราะ URI ทั้งชุดเปรียบเสมือนชื่อและตัวบ่งชี้เพื่อเข้าถึงเอกสารนั้น ถ้ามีชื่อที่ซ้ำซ้อนกันจะไม่สามารถเข้าถึงเอกสารดังกล่าวได้

1.3) อธิบายเอกสารด้วยโครงสร้างอาร์ดีเอฟ (Resource Description Framework: RDF) คุณสมบัติของเอกสารอาร์ดีเอฟ เป็นข้อมูลที่มีลักษณะเป็นโหนดและเส้น (Node and Graph) มีมุมมองในรูปแบบของแผนภาพต้นไม้ (Tree) สามารถกำหนดรากและรายละเอียดของกิ่งก้านตามระดับความสัมพันธ์ของข้อมูลได้ อาร์ดีเอฟจะใช้โครงสร้างเป็นแท็ก (Tag Element) เพื่อช่วยให้ผู้ใช้กำหนดโครงสร้างของเอกสารได้ตามความต้องการ (Klyne and Caroll, 2004) ข้อมูลที่อาร์ดีเอฟสามารถอธิบายได้จะเป็นข้อมูลที่มีจากแหล่งเดียวหรือหลายแหล่งได้ เนื่องจากมีการกำหนดแท็กเพื่อเชื่อมโยงแหล่งที่มาหรือเชื่อมโยงไปยูอาร์ไออื่นได้ ดังรูปที่ 2.10

```
<http://biglynx.co.uk/people/dave-smith>
    rdf:type foaf:Person ;
    foaf:name "Dave Smith" ;
    foaf:based_near <http://sws.geonames.org3333125/> ;
    foaf:based_near <http://dbpedia.org/resource/Wildlife_photography> ;
    foaf:topic_interest <http://dbpedia.org/resource/David_attenborough> ;
    foaf:know <http://biglynx.co.k/people/matt-briggs> ;
    rel:employerOf <http://biglynx.co.uk/people/dave-smith.rdf> .

<http://biglynx.co.uk/people/dave-smith.rdf>
    foaf:primaryTopic <http://biglynx.co.uk/people/dave-smith> .

<httP://biglynx.co.uk/people/matt-briggs>
    rel:employedBy <http://biglynx.co.uk/people/dave-smith> .
```

รูปที่ 2.10 รายละเอียดของไฟล์อาร์ดีเอฟที่มีการอ้างถึงข้อมูลยูอาร์ไอ

ข้อมูลภายในอาร์ดีเอฟจะเรียกว่าทริปเปิล (Triple) โครงสร้างของทริปเปิลแบ่งออกได้ 3 ส่วนได้แก่ ส่วนประธาน (Subject) ภาคแสดง (Predicate) และกรรม (Object) ซึ่งข้อมูลที่ประกอบเป็นเอกสารอาร์ดีเอฟ จะประกอบด้วย 2 ส่วนได้แก่ ข้อมูลเพื่ออธิบายลักษณะและคุณสมบัติของเอกสารนั้น (Literal Triples) และข้อมูลที่ใส่รายละเอียดของยูอาร์ไออื่นเพื่อทำการลิงค์ไปยังข้อมูลที่เกี่ยวข้อง (RDF Link) เพื่อใช้อธิบายความสัมพันธ์ระหว่างเอกสารทั้งสองรายการที่เชื่อมโยงกันอยู่ โดยรูปแบบของการเชื่อมโยงมี 3 ลักษณะ ได้แก่

1.3.1) การเชื่อมโยงความสัมพันธ์ (Relationship Link) เพื่อใช้เชื่อมโยงไปยังข้อมูลอื่นที่เกี่ยวข้อง เช่นเอกสารเกี่ยวกับบุคคล การเชื่อมโยงความสัมพันธ์คือการเชื่อมโยงไปยังข้อมูลอื่น ๆ ของบุคคลนั้นเช่นประวัติการทำงาน ประวัติการศึกษา ผลงานที่เคยปฏิบัติงานไว้ เป็นต้น

1.3.2) การเชื่อมโยงเพื่อบ่งชี้ (Identity Link) เพื่อใช้เชื่อมไปยังยูอาร์ไอชื่ออื่น (Aliases) ที่อยู่ในแหล่งข้อมูลอื่นที่มีข้อมูลประเภทเดียวกันอยู่ เช่น การค้นคืนข้อมูลบุคคลจากหลาย ๆ แหล่งข้อมูลเพื่อความถูกต้องของข้อมูล

1.3.3) การเชื่อมโยงคำศัพท์ (Vocabulary Link) ใช้เพื่อการเชื่อมโยงคำศัพท์ที่เกี่ยวข้องกับคำหลัก หรือการหาความหมายในหลาย ๆ มิติ หรือใช้เพื่อเชื่อมโยงไปยังเรื่องราวอื่นที่เกี่ยวข้องในมุมมองอื่น ผ่านคำศัพท์ที่น่าสนใจได้ เช่น ชื่อภาพยนตร์มีบางเรื่องที่มีชื่อเดียวกัน แต่เป็นคนละเนื้อเรื่อง ผู้ใช้สามารถเข้าถึงผ่านกลุ่มคำศัพท์ชุดเดียวกันได้

การเชื่อมโยงข้อมูลด้วยยูอาร์ไอในเครือข่ายลิงค์โอเพนดาทาจะเป็นการนำข้อมูลปลายทางที่กำหนดในยูอาร์ไอถูกนำไปใช้ต่อได้ โดยไม่ต้องสร้างข้อมูลชุดเดิมซ้ำอีก ซึ่งสามารถอ้างถึงได้ทั้งการชี้ข้อมูลที่เกี่ยวข้องอื่น ๆ โดยชี้ออกไปยังแหล่งจัดเก็บหรือยูอาร์ไออื่น ๆ (Forward Link) และการนำข้อมูลจากยูอาร์ไอื่น ๆ ที่เกี่ยวข้องกับเอกสารหลักดึงเข้ามาจัดเก็บ (Backward Link) ให้ข้อมูลในเอกสารต้นทางมีความสมบูรณ์ครบถ้วนจากแหล่งข้อมูลภายนอกได้ เช่น ข้อมูลผู้แต่ง ผู้วาดภาพ บรรณาธิการ หรือข้อมูลประกอบเพื่อความน่าเชื่อถือ รวมถึงข้อมูลที่อธิบายตัวเอกสารเอง เช่น ขนาดไฟล์ รูปแบบ ชนิดของเอกสาร ประเภทของข้อมูล ขอบเขตการนำไปใช้ หรือแม้แต่ลิขสิทธิ์หรือผู้มีสิทธิ์ในผลงานนั้น เป็นต้น

2.3.4 การผสานเมทาดาทา (Metadata Mapping)

การผสานเมทาดาทา หรือการวิเคราะห์โครงสร้างของเมทาดาทา เป็นแนวคิดที่พัฒนาขึ้นมาเพื่อแก้ปัญหาการใช้งานข้อมูลที่มีเค้าร่างเมทาดาทาต่างกัน

Candan, Kim, Liu and Suvarna (2006) ให้ความหมายของการผสานเมทาดาทาว่า เป็นกระบวนการรวมชุดโครงสร้างเมทาดาทาสองชุดที่มีโครงสร้างที่แตกต่างกัน แต่มีข้อมูลหรือโดเมนสารสนเทศเรื่องเดียวกันหรือเกี่ยวข้องกัน โดยการรวมชุดเมทาดาทานั้นจะต้องพิจารณาถึงรายละเอียดของแหล่งเก็บข้อมูล และบล็อกเมทาดาทา (Metadata Block) ซึ่งมีองค์ประกอบ 3 ชั้น (Layer) ได้แก่ ภาษาที่ใช้อธิบายเค้าร่างเมทาดาทา (Schema Definition Language) เค้าร่างเมทาดาทา (Metadata Schema) และเมทาดาทาอินสแตนส์ (Metadata Instance) โดยระดับการผสานเมทาดาทาสามารถแบ่งเป็น 3 ระดับ ได้แก่ 1) การผสานโครงสร้างภาษา (Language mapping) 2) การผสานองค์ประกอบย่อยเมทาดาทา (Metadata Element Mapping) และ 3) การแปลงอินสแตนส์ (Instance Transformation)

การผสานเมทาดาทา จึงเป็นเทคนิคการวิเคราะห์เค้าร่างเมทาดาทา เพื่อให้ระบบสามารถนำข้อมูลที่จัดเก็บด้วยเค้าร่างเมทาดาทาต่างกันมาใช้งานร่วมกันได้ โดยสูญเสียความหมายของข้อมูลน้อยที่สุด (Interoperability)

ในช่วงสิบปีที่ผ่านมางานวิจัยหลายชิ้นได้พยายามพัฒนาเมทาดาทามาตรฐานขึ้นมาเพื่อใช้เป็นมาตรฐานกลาง ที่ทุกหน่วยงานสามารถนำไปประยุกต์ใช้ในการอธิบายทรัพยากรสารสนเทศได้ทุกรูปแบบได้ อย่างไรก็ตาม ในทางปฏิบัติและด้วยนโยบายองค์กรแต่ละทำให้ไม่สามารถใช้งานได้จริง ทั้งนี้ เนื่องจากความหลากหลายของทรัพยากรสารสนเทศและความต้องการที่เฉพาะเจาะจงขององค์กร ทำให้การปรับใช้มาตรฐานกลางทำได้ยาก อุปสรรคในการวิเคราะห์โครงสร้างเมทาดาทาได้แก่ บริบทที่เกี่ยวข้อง เนื้อหาและการใช้งานทรัพยากรสารสนเทศ ส่งผลให้งานวิจัยที่ออกมามีความเฉพาะเจาะจงในบางบริบทหรือบางกรณีเท่านั้น ทำให้ไม่สามารถกำหนดเป็นมาตรฐานหลักในการทำงานร่วมกันได้ในทุกกรณี

2.3.4.1 แนวทางในการผสานเมทาดาทา

จากการปริทัศน์งานวิจัยเกี่ยวกับการจัดการเมทาดาทาเพื่อการแลกเปลี่ยนข้อมูลระหว่างระบบสารสนเทศ จะมุ่งเน้นการพัฒนาเทคนิคเพื่อให้ระบบสารสนเทศแต่ละระบบทำงานร่วมกับเมทาดาทาต่างโครงสร้างได้อย่างอัตโนมัติ หรือหมายถึงการปรับชุดเมทาดาทาที่มีโครงสร้างแตกต่างกันให้สามารถแลกเปลี่ยนข้อมูลระหว่างกันได้ Sheth และ Klas (1998) จำแนกงานวิจัยด้านการผสานเมทาดาทาออกเป็น 3 กลุ่มคือ

1) การวิจัยเทคนิคระดับล่าง (Lower Technical Level) คือ กระบวนการสร้างชุดเมทาดาทาเพื่อใช้อธิบายสื่อสารสนเทศและทำการส่งและแลกเปลี่ยนข้อมูลระหว่างระบบสารสนเทศปลายทางที่มีชุดเมทาดาทาโครงสร้างเดียวกัน

2) การวิจัยเทคนิคระดับที่สูงขึ้น (Higher Technical Level) จะเป็นการวิจัยการทำงานร่วมกันระหว่างชุดเมทาดาทาที่มีโครงสร้างแตกต่างกัน โดยระบบที่รับชุดเมทาดาทาต่างโครงสร้างเข้ามาจากระบบอื่นจะสามารถวิเคราะห์และเชื่อมโยงข้อมูลจากเมทาดาทาที่แตกต่างกันได้

3) งานวิจัยเมทาดาทาระดับเชิงความหมาย (Very High Semantic Level) จะมุ่งเน้นการวิจัยระบบเชิงความหมาย (Semantic system) โดยมีการศึกษากระบวนการวิเคราะห์เนื้อหา (Content) และความต้องการของผู้ใช้ร่วมกัน โดยสร้างระบบสารสนเทศที่มีความสามารถเข้าใจโครงสร้างของเมทาดาทาในบริบทที่ผู้ใช้ต้องการ ทั้งการแปลความหมายที่เหมาะสมกับเนื้อหา และการเชื่อมโยงเนื้อหาที่เกี่ยวข้องได้

2.3.4.2 ระดับของการผสานเมทาดาทา

จากแนวทางการวิจัยด้านการผสานเมทาดาทา สามารถแบ่งระดับการผสานเมทาดาทาได้ 3 ระดับ ดังนี้

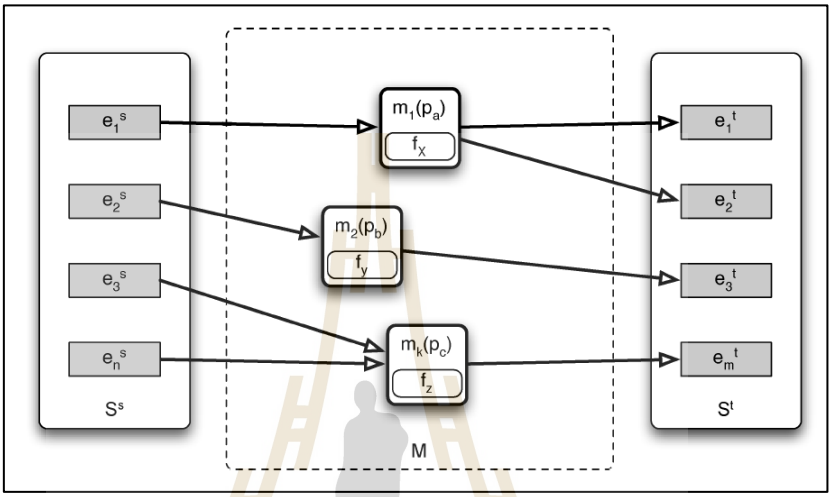
1) การผสานภาษาที่ใช้อธิบายเค้าร่างเมทาดาทา (Schema Definition Languages Mapping)

การผสานภาษาที่ใช้อธิบายเค้าร่างเมทาดาทาจะดำเนินการก็ต่อเมื่อไวยากรณ์ หรือโครงสร้างภาษาในการแสดงเมทาดาทาเป็นคนละภาษา หรือมีความแตกต่างกันโดยสิ้นเชิง ภาษาที่ใช้ในการอธิบายเค้าร่างเมทาดาทา อาทิ XML (Extensible Markup Language) OWL (Web Ontology Langue) RDF (Resource Description Framework) UML (Unified Model language) หรือ SQL-DLL (Structure Query Language Dynamic Link Library) ซึ่งภาษาเหล่านี้เป็นภาษาที่มีโครงสร้างจัดการข้อมูลคนละรูปแบบ เช่น XML เป็นภาษาแบบลำดับชั้น (Hierarchical) ส่วน UML เป็นโครงสร้างภาษาเชิงวัตถุ (Object Oriented) เป็นต้น การแปลงภาษาที่ใช้อธิบายเค้าร่างเมทาดาทานั้นอาจจะทำให้เกิดการสูญเสียรายละเอียดของข้อมูลได้ถ้าเมทาดาทานั้นมีเค้าร่างเมทาดาทาที่ต่างกัน หรือองค์ประกอบย่อยเมทาดาทาต่างกัน ดังนั้น กระบวนการผสานภาษาที่ใช้อธิบายเค้าร่างเมทาดาทา จะดำเนินการในระดับเชิงความหมาย (Very High Semantic Level)

2) การผสานองค์ประกอบย่อยเมทาดาทา (Metadata Element Mapping)

การผสานองค์ประกอบย่อยเมทาดาทา คือ การแปลงชุดขององค์ประกอบย่อยจากเค้าร่างเมทาดาต้นฉบับ ไปยังชุดองค์ประกอบย่อยของเค้าร่างเมทาดาทาเป้าหมาย โดยเค้าร่างเมทาดาทาจะประกอบด้วยองค์ประกอบย่อยตามที่กำหนดเป็นเค้าร่าง ดังนั้น ชุดขององค์ประกอบย่อยของเค้าร่างเมทาดาทาต้นฉบับ และองค์ประกอบย่อยของเค้าร่างเมทาดาทาเป้าหมายต้องมีความเกี่ยวข้องหรือเชื่อมโยงในเนื้อหาหรือคำศัพท์ที่เกี่ยวข้องกัน โดยเค้าร่างเมทาดาทาทั้งสองจะต้องมีการแปลงโครงสร้างด้วยวิธีการผสานองค์ประกอบย่อย (mapping element) ตาม

เงื่อนไขหรือกฎการผสานที่กำหนดไว้ โดยการแปลงข้อมูลจะต้องใช้คำสั่งในการผสาน (mapping expression) ซึ่งขั้นตอนการแปลงและจำนวนคำสั่งมีได้หลายรูปแบบตามลักษณะความสัมพันธ์ของเมทาดาทา ซึ่งจะต้องทำงานภายใต้ฟังก์ชั่นการทำงานรูปแบบหนึ่ง เพื่อใช้ในการกำหนดการแปลงในระดับอินสแตนส์ด้วย ดังแสดงในรูปที่ 2.11



รูปที่ 2.11 การแปลงเมทาดาทาแบบระบุองค์ประกอบย่อยของเมทาดาทาที่มีข้อมูลในกลุ่มเดียวกัน (Samir, Ioan, and Chabane, 2011)

จากรูปที่ 2.11 แสดงภาพรวมการแปลงเมทาดาทาจากเค้าร่างเมทาดาทาต้นฉบับไปยังเค้าร่างเมทาดาทาเป้าหมาย โดยผ่านกระบวนการวิเคราะห์และแปลงองค์ประกอบย่อยเมทาดาทา (Element) $m_{(k)}$ และประโยคคำสั่ง $p_{(k)}$ ซึ่งการแปลงองค์ประกอบย่อยเมทาดาทาจะมีความสัมพันธ์ในแบบ 1:1 1:N และ N:1 โดยภายในเค้าร่างเมทาดาทาทั้งสองชุดจะมีองค์ประกอบย่อยที่มีความสัมพันธ์กัน Interpretation (I) โดยความสัมพันธ์ระหว่างองค์ประกอบย่อยของเค้าร่างเมทาดาทาทั้งสองชุดจะแสดงด้วย $I(e_i^s)$ และ $I(e_j^t)$ ความเป็นไปได้ของความสัมพันธ์ระหว่างองค์ประกอบย่อยจากเค้าร่างเมทาดาทาทั้งสองชุด โดยประยุกต์ตามแนวคิดของมีนและคณะ (Mena et al., 2000) และสปาคาพิทาและคณะ (Spacapieta et al., 1992) ได้ ดังนี้

- องค์ประกอบย่อยเมทาดาทาทั้งสองไม่มีความสัมพันธ์กัน (Exclude) แสดงด้วย $I(e_i^s) \cap I(e_j^t) = \emptyset$ หมายถึงองค์ประกอบย่อยไม่มีความสัมพันธ์ระหว่างกัน หรือมีความหมายแตกต่างกันโดยสิ้นเชิง

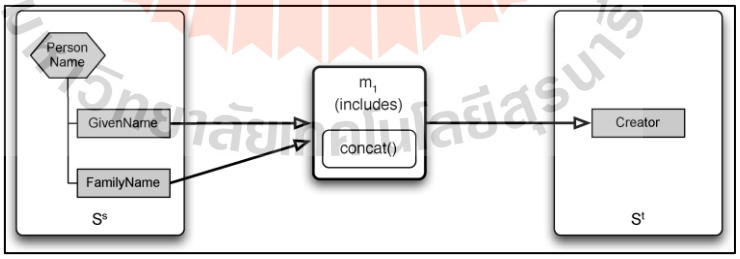
- องค์ประกอบย่อยเมทาดาทามีค่าเทียบเท่ากัน (Equavalent) $I(e_i^s) \equiv I(e_j^t)$ แสดงถึงองค์ประกอบย่อยทั้งสององค์ประกอบอาจเขียนด้วยคำศัพท์แตกต่างกัน แต่มีความหมายไปในทิศทางเดียวกัน หรือใช้แทนกันได้ เช่น Author และ Creator เป็นต้น

- องค์ประกอบย่อยเมทาดาทาเป็นส่วนย่อยขององค์ประกอบย่อยเมทาดาทาอีกตัวหนึ่ง (Include) $I(e_i^s) \subseteq I(e_j^t) \vee I(e_i^t) \subseteq I(e_j^s)$ หมายถึง องค์ประกอบย่อยของเค้าร่างเมทาดาทาต้นฉบับ หรือองค์ประกอบย่อยของเค้าร่างเมทาดาทาเป้าหมาย มีข้อมูลที่เป็นส่วนหนึ่งส่วนใดของอีกเค้าร่างเมทาดาทาหนึ่งเพียงบางส่วน เช่น FullName และ FamilyName ของเค้าร่างเมทาดาทา TVAnytme เป็นส่วนหนึ่งของข้อมูล Author หรือ Creator ในดับบลินคอร์

- องค์ประกอบย่อยเมทาดาทามีส่วนทับซ้อนกันบางส่วน (Overlap) $I(e_i^s) \cap I(e_j^t) \neq \emptyset \wedge I(e_i^s) \nsubseteq I(e_j^t) \wedge I(e_i^t) \nsubseteq I(e_j^s)$ ลักษณะความสัมพันธ์ประเภทนี้คือ องค์ประกอบย่อยหนึ่ง ๆ ไม่ได้เป็นส่วนหนึ่งส่วนใดขององค์ประกอบย่อยอื่นทั้งหมด (Not Include) แต่มีเพียงบางส่วนที่มีความเหมือนหรือเกี่ยวข้องกันเท่านั้น เช่น Description และ Info ซึ่งทั้งสององค์ประกอบย่อยนี้เก็บรายละเอียดคำอธิบายเช่นเดียวกัน แต่ Description จะมีความครบถ้วนและสมบูรณ์มากกว่า info เป็นต้น

3) การแปลงอินสแตนส์ (Instance Transformation)

การแปลงข้อมูลภายในอินสแตนส์ เป็นการปรับเปลี่ยนข้อมูลที่บันทึก (Value) ในองค์ประกอบย่อยเมทาดาทา ให้สอดคล้องและเหมะสมกับขอบเขต (Scaling) และรูปแบบการลงรายการขององค์ประกอบย่อยในเค้าร่างเมทาดาทาเป้าหมาย ดังรูปที่ 2.12 ซึ่งจะดำเนินการต่อจากการแปลงเค้าร่างเมทาดาทา



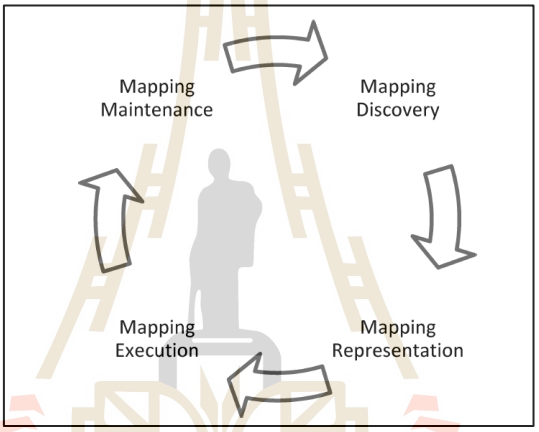
รูปที่ 2.12 การแปลงองค์ประกอบย่อยเมทาดาทาจากเค้าร่างเมทาดาทาต้นฉบับไปยังเค้าร่างเมทาดาทาเป้าหมาย (Samir, Ioan, Chabane, 2011)

จากรูปที่ 2.12 แสดงกระบวนการปรับข้อมูลจากองค์ประกอบย่อยของเค้าร่างเมทาดาทาต้นทางไปยังองค์ประกอบย่อยในเค้าร่างเมทาดาทาเป้าหมาย โดยในตัวอย่างเป็นการแปลงข้อมูลชื่อบุคคล ซึ่งเค้าร่างเมทาดาทาต้นทางมีองค์ประกอบย่อย PersonName ซึ่งประกอบด้วย

องค์ประกอบย่อยอีก 2 ตัวคือ GivenName และ FamilyName ในขณะที่เค้าร่างเมทาดาทาปลายทางที องค์ประกอบย่อยสำหรับเก็บข้อมูลบุคคลเพียงองค์ประกอบเดียวคือ Creator การนำข้อมูลจากเค้าร่างเมทาดาทาต้นทางไปใช้ในเค้าร่างเมทาดาทาปลายทางจะต้องผ่านกระบวนการแปลงข้อมูลด้วยวิธีการรวมกัน (Include) จึงจะสามารถนำข้อมูลไปเก็บในองค์ประกอบย่อยของเค้าร่างเมทาดาทาเป้าหมายคือ Creator ได้ ซึ่งกระบวนการในการแปลงข้อมูลจะใช้ฟังก์ชั่นการตัดต่อคำหรือ Concat() เพื่อรวมข้อมูลเป็นชุดเดียวกัน

2.3.4.3 กระบวนการผสานเมทาดาทา

Haslholfer (2008) ได้จำแนกกระบวนการในการผสานเมทาดาทาเป็น 4 กระบวนย่อย ดังนี้



รูปที่ 2.13 วงจรการผสานเมทาดาทา

1) การค้นหาเมทาดาทาเพื่อนำมาผสาน (Mapping Discovery) เป็นกระบวนในการวิเคราะห์ความสัมพันธ์ระหว่างระหว่างเค้าร่างเมทาดาทาต้นทางกับเมทาดาทาปลายทาง เป็นการวิเคราะห์ในเชิงความหมายเพื่อเปรียบเทียบความเหมือนขององค์ประกอบย่อยเมทาดาทาจากเค้าร่างเมทาดาทาที่ต่างกัน และทำการหารายละเอียดร่วมเพื่อเข้าสู่ขั้นตอนการปรับองค์ประกอบย่อยโดยการวิเคราะห์ถึงระดับอินสแตนส์ หรือค่าข้อมูลที่บันทึก (Value) ของแต่ละองค์ประกอบย่อย งานวิจัยของ Rahm และ Bernstein (2001) กับ Kalfoglou และ Schorlemmer (2003) ได้นำเสนอกระบวนการค้นหาเชิงความหมายเพื่อระบุความแตกต่างและความเหมือนของเค้าร่างเมทาดาทาโดยใช้เทคนิคการปรับโครงสร้างภาษาให้อยู่ในรูปแบบของกราฟ (Graph base) และใช้แผนภาพต้นไม้การตัดสินใจ (Decision Tree) ในการประเมินเค้าร่างเมทาดาทาและใช้เทคนิคการผสานออนโทโลยี (Ontology Alignment) เพื่อเชื่อมโยงความหมาย โดยการเชื่อมโยงแบ่งเป็น 2 ลักษณะ คือ การรวมออนโทโลยีเป็นออนโทโลยีใหม่ (Ontology Merging) และการเทียบและรวมกับออนโทโลยีใดออนโทโลยีหนึ่ง (Ontology Mapping)

2) การนำเสนอการผสานเมทาดาทา (Mapping Representation) เป็นขั้นตอนการประเมินและวิเคราะห์ขั้นตอนการผสานเมทาดาทา โดยการกำหนดรูปแบบการผสานเมทาดาทาตามลักษณะของเค้าร่างเมทาดาทาทั้งสอง (Noy, 2004) โดยแนวคิดในการผสานเมทาดาทาจะแบ่งเป็น 3 รูปแบบ ได้แก่ 1) โมเดลการผสานในระดับข้อมูลในอินแสตนซ์ (Instance) 2) แบบกำหนดกฎความสัมพันธ์และรูปแบบการผสานระหว่างคลาสที่พิจารณา และ 3) การเปรียบเทียบโครงสร้างคลาสและเค้าร่างเป้าหมาย

3) การดำเนินการผสานเมทาดาทา (Mapping Execution) เป็นขั้นตอนทำการระบุรายละเอียดของกระบวนการในการผสานเมทาดาทา ซึ่งการผสานสามารถใช้กระบวนการหลากหลายขึ้นอยู่กับบริบท สภาพแวดล้อมของข้อมูล และรูปแบบการทำงานของระบบที่เลือกใช้ เช่น การใช้โปรแกรมสำเร็จรูป การเขียนโปรแกรมในการแปลงเมทาดาทา การทดสอบด้วยข้อคำถาม (Query) เพื่อทดสอบผลการค้นคืนและรายละเอียดของเมทาดาทายังครบถ้วนเหมือนก่อนทำการผสานหรือไม่ (Halevy, 2001)

4) การบำรุงรักษาการผสานเมทาดาทา (Mapping Maintenance) เป็นขั้นตอนสุดท้ายในการผสานเมทาดาทา โดยเป็นขั้นตอนแบบวนซ้ำ กล่าวคือ เมื่อเมทาดาทาผสานกันแล้วจะต้องทำการทดสอบการทำงาน และติดตามเค้าร่างเมทาดาทา การติดตามหรือการตรวจสอบรุ่นของเค้าร่างเมทาดาทา (Versioning Control) เมื่อการปรับเปลี่ยนรายละเอียดของเค้าร่างเมทาดาทาอาจส่งผมต่อเมทาดาทาที่มีการผสานกันไว้แล้ว ไม่ว่าเค้าร่างต้นทางหรือเค้าร่างที่มีการผสานมีการเปลี่ยนแปลงก็ต้องมีการตรวจสอบผลการใช้งานและประเมินความเปลี่ยนแปลงที่อาจเกิดขึ้นได้ (Noy และ Musen, 2004)

ในส่วนของระบบการจัดเก็บเลิร์นนิงอ็อบเจ็กต์ งานวิจัยนี้ได้ออกแบบเค้าร่างเมทาดาทาจากการสำรวจความต้องการของผู้ใช้ โดยอ้างอิงมาตรฐานเมทาดาทาจากดับลินคอร์และลอมเมทาดาทา และทำการเลือกเขตข้อมูลที่ผู้ใช้เลือกใช้งานจริงเท่านั้น ดังนั้น การผสานเมทาดาทาจะเป็นการผสานองค์ประกอบย่อยเมทาดาทา (Metadata Element Mapping) กล่าวคือ พิจารณาเฉพาะความหมายขององค์ประกอบย่อยภายในเค้าร่างเมทาดาทาเท่านั้น โดยการพิจารณาถึงชื่อขององค์ประกอบย่อย (Element Title) ประเภทข้อมูลขององค์ประกอบย่อย (Element Data Type) เพื่อดำเนินการผสานองค์ประกอบย่อยที่เหมือนกันเข้าด้วยกัน และนำมาผสานในเค้าร่างเมทาดาทาที่กำหนดไว้ โดยไม่พิจารณาถึงโครงสร้างภาษาของเค้าร่างเมทาดาทาที่มีโดเมนต่างกัน หรือเมทาดาทาที่อธิบายข้อมูลอื่น ๆ นอกจากข้อมูลของเลิร์นนิงอ็อบเจกต์ เนื่องจากชุดข้อมูล (Dataset) ในงานวิจัยนี้จะใช้ชุดข้อมูลตัวอย่างที่เป็นเมทาดาทาที่อธิบายเลิร์นนิงอ็อบเจกต์เท่านั้น

2.4 การค้นคืนเลิร์นนิงอ็อบเจกต์จากโอเพนดาทา

ในส่วนนี้ ประกอบด้วยแนวคิดทฤษฎีที่เกี่ยวกับการค้นคืนเลิร์นนิงอ็อบเจกต์ ที่จัดเก็บในโอเพนดาทา (Open Data) ซึ่งจัดเก็บข้อมูลในโครงสร้างแบบ RDF และ OWL และทำการค้นคืนด้วยภาษาสปาเคิล (SPARQL) โดยมีรายละเอียด 4 ส่วนหลักคือ (1) ภาษาและเครื่องมือสำหรับค้นคืนโอเพนดาทา ประกอบด้วย ภาษาสปาเกิล เบราเซอร์สำหรับโอเพนดาทา เครื่องมือค้นคืนเชิงความหมาย การค้นคืนด้วย

ภาษาธรรมชาติเชิงหมายความ (2) การวัดความคล้ายคลึงของผลการค้นคืน ปะกอบด้วย การวัดความคล้ายคลึงของเอกสารด้วยปริภูมิเวกเตอร์ การวัดความคล้ายคลึงเชิงความหมาย (3) การจัดลำดับผลการค้นคืน ประกอบด้วย แนวทางในการจัดลำดับผลการค้นคืน ปัจจัยที่มีผลต่อการจัดลำดับผลการค้นคืน และ (4) บทสรุปปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.4.1 ภาษาและเครื่องมือสำหรับค้นคืนโอเพนดาทา (Open Data Searching and Tools)

2.4.1.1 ภาษาสปาเคิล (SPARQL)

สปาเคิล (SPARQL) เป็นภาษาสอบถามที่ใช้งานกับข้อมูลที่จัดเก็บในโครงสร้างของ RDF และ OWL จึงเป็นภาษาสอบถาม (Query Language) พื้นฐานที่ใช้ในการค้นคืนข้อมูลในโอเพนดาทาซึ่งจัดเก็บข้อมูลตามมาตรฐาน RDF แหลักการของการค้นคืนด้วยภาษา SPARQL แบ่งเป็น สองส่วน หลัก ๆ คือ ส่วน SELECT และ WHERE ซึ่ง SELECT เป็นการอธิบายตัวแปรที่ใช้ Prefix " ? " เป็น ตัวแปรที่เก็บค่าผลลัพธ์ และ WHERE เป็นเงื่อนไขการค้นคืนข้อมูล โดย SPARQL จะดึงข้อมูลจากไฟล์ที่ถูกจัดเก็บในรูปแบบ RDF หรือ OWL (Sbodio, 2010) ซึ่งไฟล์เหล่านี้ จะอธิบายข้อมูลในรูปแบบของกราฟที่มีส่วนประกอบ 3 ส่วน คือ Subject Predicate และ Object (Basic Graph Pattern) ดังแสดงในตารางที่ 2.2

ตารางที่ 2.2 รูปแบบโครงสร้างประโยคคำถามสปาเคิล

รูปแบบภาษา SPARQL	ตัวอย่างการใช้คำสั่ง
PREFIX	PREFIX vcard: <http://www.w3c.org/2001/vcard-rdf/3.0#>
SELECT ?varnamelist	SELECT ?x ?givenName
WHERE {basic graph pattern}	WHERE { ?x vcard:Family "Smith"} ?x vcard:Given "givenName.}

จากตารางที่ 2.2 เป็นโครงสร้างภาษาสปาเคิลเพื่อใช้ค้นหาข้อมูล givenName ของ x โดยกำหนดงื่อนไข family เป็น Smith และแสดงข้อมูลของ givenName เป็นผลลัพธ์ ทั้งนี้ข้อมูลที่จัดเก็บตามมาตรฐานอาร์ดีเอฟ ที่อยู่ในโอเพนดาทา ผู้ใช้สามารถเข้าถึงข้อมูลจากโหนด (Node) หนึ่งและข้ามไปยังโหนดอื่น ๆ ภายในกลุ่มข้อมูลเหล่านี้เรียกว่า ชุดข้อมูล (Dataset) ที่มีข้อมูลที่สัมพันธ์กันในมิติใดมิติหนึ่ง ความสัมพันธ์ของข้อมูลเหล่านี้ช่วยให้ผู้ใช้สามารถควบคุมทิศทาง และเข้าถึงข้อมูลระหว่างชุดข้อมูลได้อย่างอิสระตราบใดที่ชุดข้อมูลเหล่านั้นมีข้อมูลหรือเนื้อหาที่เกี่ยวข้องกัน ทั้งนี้ ภาษาสปาเคิลได้ถูกประยุกต์และพัฒนาเป็นเครื่องมือการเข้าถึงข้อมูลในโอเพนดาทาเพื่อให้มีความสะดวกมากกว่าการใช้ภาษาสปาเคิลโดยตรง ดังนี้

2.4.1.2 เบราเซอร์สำหรับโอเพนดาทา (Open Data Browser)

แหล่งข้อมูลในโอเพนดาทาสามารถเข้าถึงและนำมาใช้ประโยชน์ได้ผ่านทางเทคโนโลยีอินเทอร์เน็ต ผ่านทางเบราเซอร์ (Browser) แต่เบราเซอร์ทั่วไปนั้นจะรองรับมาตรฐานการแสดงผลในรูปแบบแท็ก (Tag) ของภาษาเอชทีเอ็มแอล (HTML) เป็นหลัก ซึ่งแท็กในรูปแบบอื่น ๆ อย่าง

เอ็กเอ็มแอล (XML) ซึ่งนำไปใช้เป็นโครงสร้างการจัดเก็บข้อมูลตามมาตรฐานอาร์ดีเอฟ (RDF) และอาว (OWL) จะต้องมีชุดคำสั่งในการนำพาข้อมูล (Parser) มาแสดงผลในรูปแบบที่กำหนดไว้เสียก่อน เช่น เอ็กเอ็มแอลสไตล์ชีท (XML StyleSheet) เพื่อเป็นการกำหนดรูปแบบการแสดงผลให้เข้าใจได้ง่าย

การใช้งานโอเพนดาทาผ่านเว็บจะขึ้นอยู่กับความสามารถในการออกแบบส่วนต่อประสานของเบราเซอร์ให้เหมาะกับผู้ใช้แต่ละกลุ่ม หรือการใช้งานแต่ละรูปแบบ ชไนเดอร์แมนและคณะ (Shneidernam et al, 2009) ได้แนะนำให้แบ่งระดับของผู้ใช้ออกเป็น 3 กลุ่มตามความชำนาญในสายวิชาชีพ และความสามารถในการใช้เทคโนโลยี ประกอบด้วย ผู้ใช้ทั่วไป ผู้ใช้เชิงเทคนิค และผู้ใช้ที่เชี่ยวชาญในสายงานนั้น ๆ ถึงแม้จะแบ่งกลุ่มผู้ใช้ออกเป็นกลุ่มย่อย เพื่อให้สามารถสร้างส่วนต่อประสานกับผู้ใช้อย่างเหมาะสมและปรับเปลี่ยนได้ อย่างไรก็ดี ด้วยลักษณะโครงสร้างพื้นฐานของโอเพนดาทาเองที่ข้อมูลมีปริมาณมหาศาลได้ถูกเชื่อมโยงกันไว้หมด ดังนั้นการนำเสนอข้อมูลเพียงบางส่วนหรือเฉพาะส่วนที่เกี่ยวข้องนั้นแทบเป็นไปได้ยาก เนื่องจากข้อมูลสามารถเปลี่ยนแปลงและนำเสนอชุดข้อมูล (Dataset) อื่น ๆ อย่างต่อเนื่อง จึงเป็นความท้าทายสำหรับนักวิเคราะห์ที่จะจำกัดกลุ่มข้อมูลที่จะนำเสนอและคงความสามารถในการปรับเปลี่ยนรายละเอียดของข้อมูลไปยังโดเมนอื่น ๆ ตามที่ผู้ใช้ต้องการได้ (Alahmari, Thom, Magee and Wong, 2012) นักวิจัยกลุ่มนี้มีความเห็นว่า ปัญหาการนำเสนอข้อมูลผ่านเบราเซอร์ที่พัฒนาขึ้นมาเฉพาะนั้น จำเป็นจะต้องพิจารณาถึงความต้องการของผู้ใช้ ประกอบกับขอบเขตของข้อมูลที่จะนำเสนอ โดยนำเสนอแนวทางการพัฒนาและประเด็นปัญหาในการทำงานของโอเพนดาทาไว้ ดังนี้

1) การเข้าถึงเครือข่ายลิงค์ดาทา (Linked Data Exploration) เมื่อเว็บแห่งข้อมูล (Web of Data) เป็นเครือข่ายข้อมูลขนาดใหญ่ที่เกิดจากแหล่งข้อมูลบนเว็บหลาย ๆ แหล่งเชื่อมโยงกัน จะทำให้เกิดปัญหาสองประการ คือ 1) เบราเซอร์จะนำเสนอข้อมูลที่มีขอบเขตกว้างขวางให้กับผู้ใช้ได้อย่างไรโดยที่ผู้ใช้ยังรู้สึกว่าสามารถควบคุมการทำงานได้ ไม่หลงทางในการเลือกใช้ข้อมูล และยังสามารถปรับแต่งการนำเสนอและกำหนดรูปแบบข้อมูลที่ต้องการได้ และ 2) เบราเซอร์จะสามารถเชื่อมโยงข้อมูลในชุดข้อมูล (Dataset) ต่าง ๆ และสามารถเชื่อมโยงข้อมูลบนเว็บไซต์ทั่วไปที่ไม่ได้มีโครงสร้างข้อมูลเหมือนลิงค์โอเพนดาทาไปพร้อม ๆ กันได้หรือไม่ เนื่องจากผู้ใช้คงไม่ต้องการใช้งานเบราเซอร์หลาย ๆ ตัวเพื่อเข้าถึงข้อมูลพร้อมกัน

2) การควบคุมทิศทาง (Navigation) การควบคุมทิศทางในการใช้เบราเซอร์สำหรับเครือข่ายโอเพนดาทาจะมีความแตกต่างจากการใช้งานเบราเซอร์ทั่วไป กล่าวคือ เบราเซอร์ทั่วไปจะใช้การสร้างไฮเปอร์ลิงค์ (Hyper Link) เพื่อเชื่อมโยงไปยังข้อมูลหรือหน้าเอกสารอื่น ๆ ได้ ในขณะที่เบราเซอร์สำหรับโอเพนดาทาจะต้องใช้การเชื่อมโยงผ่านโครงสร้างภาษาอาร์ดีเอฟ (RDF) และจัดเก็บข้อมูลในรูปแบบของยูอาร์ไอ (URI) ทำให้การเข้าถึงข้อมูลมีข้อจำกัดในการเข้าถึงข้อมูลตามเมทาดาทา ไม่สามารถเชื่อมโยงตามข้อความเหมือนการค้นคืนข้อมูลจากเบราเซอร์ทั่วไปได้

3) การปฏิสัมพันธ์ (Interactivity) การใช้งานข้อมูลในโอเพนดาทาขึ้นอยู่กับการปฏิสัมพันธ์ของผู้ใช้กับระบบในการเลือกชุดข้อมูล และองค์ประกอบย่อยเมทาดาทา รวมถึงมุมมองข้อมูล (Facet) ในมิติต่าง ๆ อย่างไรก็ตาม การใส่ความสามารถในการปรับแต่งหรือมีปฏิสัมพันธ์จำนวนมากจะเกิดความยุ่งยากในการจัดการโครงสร้างข้อมูลที่มีภาษาแตกต่างกัน (RDF, OWL และ SPARQL) การพัฒนาก็ยิ่งยุ่งยากขึ้นเนื่องจากต้องมีการปรับปรุงฟังก์ชั่นให้รองรับกับโครงสร้างข้อมูลที่แตกต่างกัน ดังนั้น การวิจัยเพื่อสร้างระบบแสดงข้อมูลมักจะเน้นที่ลักษณะข้อมูลที่มี มากกว่าการเพิ่มความสะดวกให้กับผู้ใช้อย่างไม่มีขีดจำกัด เพราะผลกระทบจะเกิดกับผู้ดูแลระบบและส่งผลต่อการทำงานของระบบที่ไม่มีเสถียรภาพ (Stability)

ในปัจจุบันมีเบราเซอร์สำหรับแสดงข้อมูลจากโอเพนดาทาหลายตัว ซึ่งแต่ละตัวจะมีความโดดเด่นหรือฟังก์ชั่นรองรับการนำเสนอข้อมูลที่แตกต่างกัน จึงมีนักวิจัยหลายกลุ่มที่กำหนดตัวชี้วัด (Key Indicator) เพื่อใช้เป็นตัวประเมินความครบถ้วนของเบราเซอร์ที่ควรจะมีฟังก์ชั่นต่าง ๆ เพื่อรองรับการใช้งานโอเพนดาทา

งานวิจัยนี้ได้รวบรวมตัวชี้วัดสำหรับประเมินเบราเซอร์โดยพิจารณาจากความสามารถหลักที่นักวิจัยในหลายสถาบันได้นำเสนอไว้ ดังตารางที่ 2.3 (Nikolov, Uren, Motta, and Roeck, 2008; Niles and Pease, 2001)

ตารางที่ 2.3 เปรียบเทียบตัวชี้วัดเบราเซอร์สำหรับลิงค์โอเพนดาทา

ตัวชี้วัด	Disco	Marble	Sig.ma	Fenfire	Tabulator	ODE	Sparrax	Factes
Visualization				X	X	X		
RDF Graph				X		X		
Detail On Demand				X	X	X		
Highlight Link		X		X		X		
Scalability			X	X		X		
Query (Syntax)	X	X		X	X	X	X	X
Query (Keyword)			X		X	X		
Filtering			X	X		X		
History			X		X	X		
Template		X	X			X		
Entry Point		X	X			X		

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตารางที่ 2.3 เปรียบเทียบตัวชี้วัดเบราเซอร์สำหรับลิงค์โอเพนดาทา (ต่อ)

ตัวชี้วัด	Disco	Marble	Sig.ma	Fenfire	Tabulator	ODE	Sparrax	Factes
Faceted Browse							X	X
Reusable output	X		X	X	X	X		
Plugin available				X	X			
Forward Nav.		X	X	X	X	X		
Backward Nav.		X	X	X	X	X		
Export RDF/JSON			X		X	X		
Nav. Global LOD			X	X	X	X		
Nav. Local LOD			X	X	X	X		

จากตารางที่ 2.3 จะเห็นได้ว่า ODE เป็นเบราเซอร์ที่มีคุณสมบัติครบถ้วนมากที่สุดจากการประเมิน อย่างไรก็ตาม Denfire, Tabulator ก็มีความสามารถในการสร้างกราฟิกวิชวล (Visualization) เพื่อแสดงข้อมูลได้เช่นเดียวกัน และยังมีความสามารถในการเพิ่มส่วนเสริม (Plug -in) เข้ากับเบราเซอร์ทั่วไปอย่างไฟฟอกซ์ (Firefox) และเบราเซอร์อื่น ๆ ได้ ส่วน Sparrax และ Factes มีความสามารถในการไล่เรียงผลข้อมูล (Facet Browse) และปรับแต่งการแสดงผลในส่วนย่อย (Sub Selection) และสามารถใช้รูปประโยคคำถาม SPARQL ในการค้นคืนผ่าน SPARQL Endpoint ได้โดยตรง อย่างไรก็ดีการใช้ประโยคคำถามโดยตรงอาจมีปัญหาเรื่องเวลาในการตอบสนอง (Response Time) ถ้าชุดข้อมูลมีขนาดใหญ่ เนื่องจากใช้การเข้าถึงแบบเรียงลำดับ ส่วนเบราเซอร์ที่ดีอาจมีฟังก์ชั่นที่เพิ่มความสะดวกให้กับผู้ใช้ เช่น การใช้ผลการค้นคืนซ้ำได้ หรือการนำข้อมูลออก (Export) ในรูปแบบที่นำไปใช้ต่อได้เช่น RDF หรือ JSON ก็เป็นส่วนเสริมที่ทำให้ระบบมีการทำงานร่วมกับระบบอื่นได้ง่ายขึ้น

จากตัวชี้วัดข้างต้น ผู้วิจัยได้นำแนวคิดดังกล่าวมาใช้ในการออกแบบและพัฒนาโมเดลระบบการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ให้มีความสามารถตามตัวชี้วัดเพื่อรองรับการใช้งานที่เหมาะสมสอดคล้องกับความต้องการของผู้ใช้งาน และเหมาะสมกับการทำงานของระบบที่พึงมี

2.4.1.3 เครื่องมือค้นคืนเชิงความหมาย (Semantic Search Tool)

โปรแกรมค้นหาหรือเสิร์ชเอนจิน (Search Engine) ทำหน้าที่เก็บข้อมูลหรือเว็บไซต์ในเครือข่ายอินเทอร์เน็ตมาวิเคราะห์และจัดเก็บลงในระบบ เพื่อให้บริการการค้นคืนแก่ผู้ใช้

เช่นเดียวกับเครือข่ายชุดข้อมูลในลิงค์โอเพนดาทาที่มีเครื่องมือในการเข้าถึงนอกเหนือจากเบราเซอร์สำหรับลิงค์โอเพนดาทา ที่สามารถเข้าถึงข้อมูลทางตรง แสดงผลตามรายละเอียดของชุดข้อมูลที่จัดเก็บหรือเชื่อมโยงไปยังเครือข่ายลิงค์โอเพนดาทาภายนอก ยังมีเครื่องมือค้นคืนเชิงความหมายหรือเรียกอีกชื่อหนึ่งว่า เสิร์ชเอนจินเชิงความหมาย (Semantic Search Engine) เป็นเครื่องมือช่วยค้นคืนข้อมูลจากชุดข้อมูลผ่านคำค้นของผู้ใช้ โดยที่ผู้ใช้ไม่จำเป็นต้องทราบรายละเอียดของโครงสร้างข้อมูล ภาษาในการสอบถาม หรือหลักไวยากรณ์ต่าง ๆ ในการติดต่อกับชุดข้อมูล

เครื่องมือช่วยค้นคืนเชิงความหมายเหล่านี้มีการพัฒนาขึ้นหลายชื่อที่ได้รับความนิยม เช่น Falcons, SWSE (Semantic Web Search Engine) Sidice Swoogle และ Watson ซึ่งภายในระบบค้นคืนเหล่านี้จะทำหน้าที่รับคำค้นจากผู้ใช้และทำการดึงข้อมูลของชุดข้อมูลแต่ละชุดเพื่อเปรียบเทียบรายละเอียดที่เกี่ยวข้องกับคำค้นเพื่อนำมาแสดงผล โดยองค์ประกอบในการทำงานของเครื่องมือค้นคืนในลิงค์โอเพนดาทาได้เปรียบเทียบและแสดงในตารางที่ 2.4

ตารางที่ 2.4 เปรียบเทียบฟังก์ชั่นของเครื่องมือค้นคืนเชิงความหมายสำหรับลิงค์โอเพนดาทา

ชื่อเครื่องมือ	ส่วนต่อประสาน	ผลการค้นคืน	ประเภทเอกสาร	หมายเหตุ
SWSE	ส่วนต่อประสานออกแบบตามการใช้งานของมนุษย์	แสดงลิงค์ โครงสร้างข้อมูล จำนวนเอนทิตีที่เกี่ยวข้อง จำนวนเอนทิตีที่เลือกมาแสดง	RDF/XML ร่วมกับ HTML และ CSS เพื่อปรับการแสดงผล	SWSE สามารถลำดับการแสดงผลตามเอนทิตีของโครงสร้างข้อมูลได้
Falcon	ส่วนต่อประสานออกแบบตามการใช้งานของมนุษย์	แสดงลิงค์ โครงสร้างข้อมูล จำนวนเอนทิตีที่เกี่ยวข้อง จำนวนเอนทิตีที่เลือกมาแสดง	โครงสร้าง RDF/XML เท่านั้น	มีชุดข้อมูลกว่า 30 ล้านชุดที่อยู่ในรายการดัชนี
SinDice	ส่วนต่อประสานออกแบบตามการเชื่อมต่อกับเว็บเซอร์วิส	RDF/XML/JSON/HTML/Plain Text	RDF, Web	รองรับมาตรฐานการนำข้อมูลออกหลายมาตรฐาน แต่การทำงานค่อนข้างช้ากว่าระบบอื่น ๆ

ตารางที่ 2.4 เปรียบเทียบฟังก์ชั่นของเครื่องมือค้นคืนเชิงความหมายสำหรับลิงค์โอเพนดาทา (ต่อ)

ชื่อเครื่องมือ	ส่วนต่อประสาน	ผลการค้นคืน	ประเภทเอกสาร	หมายเหตุ
Swoogle	ส่วนต่อประสานออกแบบตามการเชื่อมต่อกับเว็บเซอร์วิส	Ontology (OWL + RDF file)	OWL, RDF ร่วมกับ HTML เพื่อการแสดงผล	มีรายการในดัชนีกว่า 1.3 ล้านรายการซึ่งมีจำนวนข้อมูลกว่า 240 ล้านทริปเปิล ใช้การเรียงลำดับผลการค้นด้วย Page Rank ของกูเกิล
Watson	ส่วนต่อประสานออกแบบตามการเชื่อมต่อกับเว็บเซอร์วิส	RDF/XML	เอกสาร RDF	รองรับการใช้คำสำคัญ และการใช้ SPARQL ในการค้นคืนข้อมูล

จากตารางที่ 2.4 จะเห็นได้ว่าเครื่องมือการค้นคืนมีสองลักษณะคือ ออกแบบมาเพื่อให้ผู้ใช้ที่เป็นมนุษย์หรือมีการแสดงผลการค้นคืนที่คนเข้าใจได้ อีกกลุ่มเป็นการออกแบบเพื่อสื่อสารกับเว็บเซอร์วิสระบบอื่นเพื่อส่งข้อมูลหรือแลกเปลี่ยนข้อมูล ทั้งนี้ระบบการค้นคืนมีพื้นฐานการจัดเก็บหรือดึงข้อมูลในรูปแบบอาร์ดีเอฟและโอดับเบิ้ลยูแอล ที่เหมือนกัน แต่ภายในมีสถาปัตยกรรมและอัลกอริทึมการเปรียบเทียบ วิเคราะห์ข้อมูลที่แตกต่างกันไป (Dietze, and Schroeder, 2009) กลุ่มนักวิจัยได้นำเสนอตัวชี้วัดเพื่อประเมินเครื่องมือค้นคืน ตัวชี้วัดประเมินถึง 1) โครงสร้างออนโทโลยี 2) ประเภทของเอกสารที่ค้นคืนหรือเข้าถึงได้ 3) การแบ่งกลุ่มผลการค้นคืนหรือระดับการเข้าถึงสามารถแบ่งกลุ่มการนำเสนอ การบังคับทิศทาง หรือการกำหนดชุดข้อมูลที่ต้องการ และ 4) จำนวนของทริปเปิลที่ค้นพบ โดยปัจจัยเหล่านี้ขึ้นอยู่กับโดเมนของเนื้อหาและโครงสร้างข้อมูลที่ใช้ในการจัดเก็บข้อมูล ซึ่งเมื่อนำเครื่องมือค้นคืนข้างต้นมาประเมินแล้ว ทุกรายการมีองค์ประกอบเดียวกันหมดคือ ใช้โครงสร้างการจัดเก็บข้อมูลเป็นเอกสารอาร์ดีเอฟ และสร้างเป็นออนโทโลยีเพื่อสามารถเข้าถึงและค้นคืน เชิงความหมายได้ และเอกสารหรือผลการค้นคืนจะนำเสนอเป็นเอกสารอาร์ดีเอฟเป็นพื้นฐาน แต่จะมีเครื่องมือค้นคืนที่ออกแบบตามการใช้งานของผู้ใช้จะมีการแปลงการแสดงผลด้วยคำสั่งสคริปต์อีกชั้นหนึ่ง เพื่อปรับรูปแบบการนำเสนอได้แก่

SWSE และ Falcon ส่วน Sindice Swoogle และ Watson เป็นระบบค้นคืนที่เน้นบริการให้กับเว็บเซอร์วิสอื่น ๆ จึงไม่ได้พัฒนาส่วนประสานกับผู้ใช้มากนัก แต่ใช้การแสดงผลเป็นเอกสารอาร์ดีเอฟโดยตรง

2.4.1.4 การค้นคืนเชิงความหมาย (Natural Language Semantic Search)

เป็นกระบวนการค้นคืนที่ช่วยทำให้คอมพิวเตอร์เข้าใจภาษาธรรมชาติของมนุษย์ ขบวนการของการประมวลผลภาษาธรรมชาติเริ่มต้นจากการรับเข้าข้อมูลที่เป็นข้อความภาษาธรรมชาติ จากนั้นระบบจะวิเคราะห์ข้อมูลทั้งในเชิงโครงสร้างและความหมายทางภาษา โดยทำงานร่วมกับฐานความรู้แล้วประเมินค่าเพื่อหาคำตอบ ทำการแทนความหมายและสร้างผลลัพธ์ที่เป็นภาษาธรรมชาติออกมา โดยมีองค์ประกอบต่าง ๆ ดังนี้

1) การวิเคราะห์ในเชิงโครงสร้าง (Syntactic Analysis) เป็นการตรวจสอบโครงสร้างทางไวยากรณ์เกี่ยวกับการวางตำแหน่งของกลุ่มคำประเภทต่าง ๆ ที่รวมกันเป็นประโยค เช่น ประโยค “The boys very fast run” เป็นประโยคที่มีโครงสร้างผิดหลักไวยากรณ์ เนื่องจากลำดับที่ถูกจะต้องเป็น “The boys run very fast”

2) การวิเคราะห์ในเชิงความหมาย (Semantic Analysis) เป็นการตรวจสอบความถูกต้องในเชิงความหมายของประโยค เช่น “The paper cut the wood” โครงสร้างของประโยคถูกต้องตามหลักไวยากรณ์ แต่เมื่อวิเคราะห์ในเชิงความหมายพบว่า มีความหมายที่เป็นไปไม่ได

3) การวิเคราะห์ในเชิงตีความ (Pragmatic Analysis) เป็นการตีความตามสถานการณ์ที่เกิดขึ้น โดยทั้งผู้ส่งข่าวสารและผู้รับข่าวสารจะต้องอยู่ในสถานการณ์เดียวกัน เช่น ขณะยืนอยู่ที่สนามบินและกำลังกังวลว่าขณะนี้เวลาเท่าไร หากถามว่า “Do you have a watch?” และได้คำตอบ “yes” หรือ “no” แสดงว่าเป็นคำตอบที่ผิด เพราะคำตอบที่ต้องการคือเวลา ณ ขณะนั้น การประมวลผลภาษาธรรมชาติ ประกอบด้วยขั้นตอนสำคัญ คือ การตัดคำและกำกับหน้าที่ของคำ (Word Segmentation and Part-of Speech Tagging) การวิเคราะห์นิพจน์ระบุนาม (Name Entity Analysis) และการวิเคราะห์นามวลี (Phrase Analysis)

ในโครงสร้างการจัดเก็บข้อมูล ลิงค์โอเพนดาทาจะรองรับชุดคำสั่งส่วนต่อประสานภาษาธรรมชาติ (Natural Language Interfaces: NLIs) ผู้ใช้สามารถสร้างประโยคคำถามด้วยภาษามนุษย์เพื่อค้นคืนข้อมูลผ่านระบบได้โดยไม่ต้องพึ่งพาความรู้โครงสร้างภาษาอาร์ดีเอฟ (RDF), สปาร์เคิล (SPARQL) เลย จากงานวิจัยของ Kaufmann and Bernstein (2007) ได้อภิปรายถึงส่วนต่อประสานสำหรับคำถามภาษาธรรมชาติ 4 รูปแบบ คือ Ginseng, NLP-Reduce, Querix และ Semantic Crystal ซึ่ง NLI เหล่านี้สามารถแปลงคำถามในรูปแบบภาษาธรรมชาติในเป็นโครงสร้างประโยคคำถาม SPARQL

รูปประโยคคำถามภาษาธรรมชาติจะประกอบด้วยคำสำคัญและองค์ประกอบของโครงสร้างภาษา ซึ่งอาจมีคำเชื่อมหรือคำขยายที่ทำให้รูปแบบประโยคมีความซับซ้อน แต่เมื่อผ่านกระบวนการแปลงเป็น SPARQL โดยการใช้ชุดคำสั่งเอพีไอ (Application Programming Interface: API) ของเจนา (Jena Framework) ตามข้อมูลในตารางที่ 2.5 แสดงข้อมูลสรุปรายละเอียดฟังก์ชั่นของ NLI แต่ละรายการ (Kaufmann and Bernstein, 2007)

ตารางที่ 2.5 ส่วนต่อประสานภาษาธรรมชาติสำหรับการค้นคืนเชิงความหมาย

NILs	สนับสนุนรูปแบบการค้นคืน	เทคโนโลยีที่ใช้
NLP-Reduce	คำสำคัญ/รูปประโยคภาษาธรรมชาติ	SPARQL + Jena + Pellet Reasoner + OWL + Knowledge Base
Qurix	รูปประโยคภาษาธรรมชาติ	SPARQL + Jena + WordNet + OWL + Knowledge Base
Ginseng	รูปประโยคภาษาธรรมชาติ	SPARQL + Jena + OWL + Knowledge Base

จากตารางที่ 2.5 ส่วนต่อประสานภาษาธรรมชาติสำหรับการค้นคืนข้อมูลในโอเพนดาทาสามารถทำให้ทั้งในรูปแบบคำสำคัญ และรูปประโยคภาษาอังกฤษ ถ้าพิจารณาถึงเทคโนโลยีที่จำเป็นต่อการพัฒนาระบบค้นคืนภาษาธรรมชาติเชิงความหมายที่ทุก NLI มีคือ API ของเจนาเฟรมเวิร์ค (Jena Framework) เพื่อใช้เป็นชุดคำสั่งในการแปลงประโยคคำถามให้เป็นโครงสร้างประโยค SPARQL ทั้งนี้ยังต้องมีส่วนของฐานความรู้ (Knowledge base) เพื่อทำหน้าที่เก็บรูปประโยคคำถาม และความรู้ที่ใช้เป็นคำตอบ

ทั้งนี้โมเดลในงานวิจัยนี้ไม่ได้ครอบคลุมส่วนคำสั่งการค้นคืนด้วยภาษาธรรมชาติ แต่มีส่วนขยาย (Extension) ที่สามารถพัฒนาฐานความรู้และนำมาเชื่อมต่อกับโมเดลเพื่อรองรับการค้นคืนด้วยภาษาธรรมชาติได้ในอนาคต

2.4.1.5 ความท้าทายการค้นคืนข้อมูลในโอเพนดาทา

ถึงแม้ว่าโอเพนดาทานั้นจะมีข้อดีในเรื่องการแลกเปลี่ยนข้อมูล การใช้ข้อมูลที่มีอยู่แล้วร่วมกันโดยไม่ต้องสร้างขึ้นใหม่ แต่เนื่องจากจำนวนข้อมูลที่มีปริมาณมากและเชื่อมโยงกันหมด ทำให้เครื่องมือค้นคืนในโอเพนดาทายังมีปัญหาในเรื่องความแม่นยำ (Precision) ค่อนข้างต่ำและได้จำนวนผลการค้นคืน (Recall) ที่สูงเกินไป ดังนั้น จึงต้องหาแนวทางในการเพิ่มความแม่นยำ

และลดจำนวนผลลัพธ์ให้เหลือรายการที่ถูกต้องหรือเกี่ยวข้องเท่านั้น นอกจากเรื่องความถูกต้องของผลการค้นคืนแล้ว ปัญหาในการค้นคืนผ่านชุดข้อมูล (Dataset) ยังมีโอกาสที่จะค้นหาผิดชุดข้อมูลหรือไปค้นหาในโดเมนที่ไม่เกี่ยวข้องกับหัวเรื่องที่ต้องการ ซึ่งส่งผลให้ใช้เวลาในการค้นคืนนานกว่าการค้นจากระบบฐานข้อมูลทั่วไปที่มีเรื่องหรือโดเมนเท่าที่กำหนดไว้ ทั้งนี้ขึ้นอยู่ผู้ใช้ที่รู้ขอบเขตเรื่องที่ต้องการก็สามารถกำหนดชุดคำถามที่เหมาะสมได้เช่นกัน

จากปัญหาข้างต้น ความถูกต้องของการค้นคืนในโอเพนดาทาจึงต้องใช้ทั้งกระบวนการเชิงความหมายและเชิงไวยากรณ์ของรูปประโยคด้วย (Madhu, Govardhan, and Rajinikanth, 2011) กระบวนการในการค้นคืนด้วยภาษาธรรมชาติต้องการกระบวนการแปลงประโยคคำถามให้อยู่ในรูปแบบโครงสร้างและคำศัพท์อาร์ดีเอฟ รวมถึงการจัดการกับประเภทหมวดหมู่ของข้อมูลที่ซับซ้อน การใช้ฟังก์ชั่นเปรียบเทียบ และการใช้เทคนิคเรื่องการให้เหตุผล (Reasoning) ร่วมด้วย (Citec, 2011)

จากแนวคิดการพัฒนาเครื่องมือการค้นคืนและเข้าถึงเลิร์นนิงอ็อบเจกต์ในโอเพนดาทา จะเห็นได้ว่าโครงสร้างพื้นฐานการจัดเก็บข้อมูลด้วย RDF, OWL ในโอเพนดาทาสามารถรองรับเทคโนโลยีเชิงความหมายและระบบฐานความรู้เพื่อค้นคืนด้วยภาษาธรรมชาติได้ ทั้งนี้ในงานวิจัยชิ้นนี้ ได้นำแนวทางข้างต้นมาออกแบบระบบค้นคืนเลิร์นนิงอ็อบเจกต์ โดยพัฒนาในรูปแบบของการค้นหา (Searching) และแบบไล่เรียง (Browsing) โดยประมวลผลร่วมกับคลังคำศัพท์ทางคณิตศาสตร์ที่แสดงความสัมพันธ์ของคำศัพท์ตามโครงสร้างของธิซอรัส ในขั้นตอนการค้นคืน ทั้งนี้เพื่อให้ระบบสามารถวิเคราะห์โครงสร้างเนื้อหาและช่วยปรับปรุงคำค้นให้เหมาะสมมากขึ้น รวมถึงยังช่วยให้การกำหนดขอบเขตการค้นคืนในเรื่องที่เกี่ยวข้องตามระดับความลึกและความกว้างของหัวเรื่องที่เกี่ยวข้องให้แม่นยำมากขึ้น นอกจากนี้ ในส่วนการค้นคืนยังใช้การวัดความคล้ายคลึงเชิงความหมาย และการจัดลำดับของผลการค้นคืนซึ่งจะอธิบายถึงเทคนิคและวิธีการในลำดับถัดไป

2.4.2 การวัดความคล้ายคลึงของผลการสืบค้นค้น (Similarity Measurement)

แนวคิดในการวัดความคล้ายคลึงระหว่างเอกสารในระบบสารสนเทศโดยตรงมีอยู่หลายเทคนิควิธี Salton and McGill (1983) ได้ทำการศึกษาการวัดความคล้ายคลึงของเอกสารโดยใช้วิธีการวัดค่าความคล้ายคลึงเชิงมุม (Cosine Similarity) ซึ่งเป็นวิธีการวัดความคล้ายคลึงของโมเดลปริภูมิเวคเตอร์ (Vector Space) เพื่อวัดการทำมุมต่อกันของเอกสาร Strasberg, Manning, Rindfleisch and Melmon (2000) ได้ศึกเปรียบเทียบเทคนิควิธีต่าง ๆ ในการวัดความคล้ายคลึงของเอกสาร โดยการวัดค่าความคล้ายคลึงของเอกสารที่เลือกมาเป็นเอกสารตัวอย่างและเอกสารอื่นทั้งที่

เกี่ยวข้องและไม่เกี่ยวข้อง พบว่า ปัจจัยที่ทำให้ได้ผลต่างกันได้แก่ วิธีการเตรียมข้อมูลนำเข้า การคิดค่าถ่วงน้ำหนักของคำ นอกเหนือจากนี้จะเป็นการประยุกต์ใช้การวัดความคล้ายคลึงระหว่างเอกสารในระบบต่าง ๆ เช่น การแบ่งกลุ่มเอกสารตามการวิเคราะห์หัวเรื่องและหมวดหมู่เอกสารโดยใช้มาตรฐานการจัดหมวดหมู่และลงรายการ (Cutting, Karger, Pedersen, and Tukey, 1992; Sahami, Yusufali and Baldonado,1998) และการค้นหาเอกสารที่เกี่ยวข้องกับเอกสารที่ผู้ใช้กำหนดหัวเรื่องและขอบเขตการนำเสนอข้อมูลเอง (Shatkay, and Wibur, 2000; Bollacker, Lawrence, and Giles, 2000)

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า การวัดความคล้ายคลึงที่นิยมใช้คือ การวัดความคล้ายคลึงเชิงมุม (Cosine Similarity) ซึ่งเป็นการแทนเอกสารด้วยระบบเวกเตอร์และวัดมุมที่กระทำต่อกันระหว่างเอกสาร เอกสารที่มีความคล้ายคลึงกันมากจะทำมุมระหว่างกันน้อย แต่อย่างไรก็ดี วิธีการที่พบในอดีตจะเน้นการประมวลผลเชิงคำนวณในการวัดระยะห่าง และการคิดค่าตัวเลขที่ได้จากเอกสารโดยตรง อีกวิธีหนึ่งคือ การวัดความคล้ายคลึงเชิงความหมายซึ่งช่วยเพิ่มประสิทธิภาพในการค้นคืนและวัดความคล้ายคลึงของเอกสารได้ดียิ่งขึ้น

จากการปริทัศน์งานวิจัยประเภทการสร้างฐานความรู้เกี่ยวกับความสัมพันธ์ของคำมีเพิ่มขึ้น และมีงานวิจัยหลายงานที่นำความรู้พื้นฐานเหล่านี้ไปประยุกต์ใช้ในด้านต่าง ๆ เช่น การนำอภิธานศัพท์ (Thesaurus) ไปช่วยปรับปรุงการสร้างกฎความสัมพันธ์ในการค้นคืน (Blake and Pratt, 2001) และนำไปประยุกต์ใช้จัดผลการค้นคืนได้ ในงานวิจัยนี้จึงได้เสนอแนวทางในการปรับปรุงประสิทธิภาพในการวัดความคล้ายคลึงระหว่างเอกสารโดยการนำความรู้พื้นฐานเกี่ยวกับความสัมพันธ์ของคำในเชิงความหมายมาใช้ รวมถึงโครงสร้างหัวเรื่องต่าง ๆ เพื่อใช้เปรียบเทียบเอกสารที่คล้ายคลึงกันกับคำค้นจากผู้ใช้ เพื่อช่วยให้ผลการค้นคืนมีความถูกต้องตามความต้องการของผู้ใช้มากที่สุด

2.4.2.1 การวัดความคล้ายคลึงของเอกสารด้วยปริภูมิเวกเตอร์

การวัดความคล้ายคลึงของเอกสารด้วยปริภูมิเวกเตอร์ (Vector space model) คือการคำนวณหาค่าความคล้าย หรือการเปรียบเทียบข้อมูล ประกอบด้วยกระบวนการย่อย 3 ขั้นตอน คือ 1) การจัดทำดัชนีให้กับเอกสารด้วยการนับค่าความถี่ของคำ 2) การกำหนดค่าน้ำหนักให้กับคำที่ดึงออกมาเป็นคำดัชนีเพื่อให้ได้เอกสารที่ตรงกับความต้องการของผู้ใช้มากขึ้น และ 3)การจัดอันดับผลการค้นที่ตรงกับคำถาม (Query) ตามค่าความคล้ายของเวกเตอร์ของเอกสาร (Document Vector) และเวกเตอร์ของคำถาม (Query Vector) โดยการวัดความคล้ายคลึงของเอกสารจะอาศัยค่าน้ำหนักของคำดัชนีที่มีอยู่ในเอกสารแต่ละรายการและค่าน้ำหนักของคำค้นที่

ผู้ใช้ใส่เข้ามาในระบบ เมื่อหาค่าน้ำหนักคำดัชนีในเอกสารและค่าน้ำหนักคำค้นจากนั้นจะนำไปแทนที่ในสมการ (ตารางที่ 2.6) เพื่อคำนวณหาค่าความคล้ายคลึงของเอกสารและคำค้น

1) เทคนิควิธีการที่ใช้ในการวัดค่าความคล้ายคลึงของเอกสาร มี 4 เทคนิควิธี คือ Inner Product, Dice, Cosine และ Jaccard

1.1) วิธีการวัดค่าความคล้ายคลึงของเอกสารแบบ Inner Product เป็นวิธีการเปรียบเทียบความคล้ายคลึงของเอกสาร โดยแต่ละเอกสารจะถูกแทนด้วยเวกเตอร์ขนาดเอ็น (N-Dimensional Vector) ซึ่งเก็บค่าน้ำหนักคำแต่ละคำในเอกสารนั้น (N-Dimensional Vector in Term Space) การเปรียบเทียบความคล้ายคลึงของเอกสารจะเปรียบเทียบโดยหาผลรวมของผลการคูณระหว่าง 2 เวกเตอร์ของเอกสาร หากเอกสารทั้งสองเอกสารคล้ายคลึงกันมาก ค่าผลรวมของผลการคูณระหว่าง 2 เวกเตอร์ของเอกสาร จะมีค่ามาก ถ้าเอกสารทั้งสองเอกสารคล้ายคลึงกันน้อย ค่าผลรวมของผลการคูณระหว่าง 2 เวกเตอร์ของเอกสารจะมีค่าน้อย

$$\sum_{k=1}^{n} w_{ik} \cdot w_{jk}$$

เมื่อ *Wi* คือ ค่าน้ำหนักของคำสำคัญจากชุดคำค้น

Wj คือ ค่าน้ำหนักของคำสำคัญในเอกสารเปรียบเทียบ

n คือ จำนวนคำสำคัญที่ทำการเปรียบเทียบ

1.2) วิธีการวัดค่าความคล้ายคลึงของเอกสารแบบ Dice เป็น วิธีการเปรียบเทียบความคล้ายคลึงของเอกสารสองเอกสาร โดยแต่ละเอกสารจะถูกแทน ด้วยเวกเตอร์ขนาดเอ็น (N-Dimensional Vector) ซึ่งเก็บค่าน้ำหนักคำแต่ละคำในเอกสาร นั้น (N-Dimensional Vector in Term Space) การเปรียบเทียบความคล้ายคลึงของ เอกสารโดยหาผลรวมของผลการคูณระหว่าง 2 เวกเตอร์ของเอกสารคูณด้วย 2 แล้วหารด้วยผลบวกของผลรวมของแต่ละเวกเตอร์ หากเอกสารทั้งสองเอกสารคล้ายคลึงกันมาก ค่าความคล้ายคลึงจะมีค่ามาก ถ้าเอกสารทั้งสองเอกสารคล้ายคลึงกันน้อย ค่าความคล้ายคลึงจะมีค่าน้อย

$$\frac{2[\sum_{k=1}^{n} w_{ik} . w_{jk}]}{\sum_{k=1}^{n} w_{ik} + \sum_{k=1}^{n} w_{jk}}$$

เมื่อ *Wi* คือ ค่าน้ำหนักของคำสำคัญจากชุดคำค้น

Wj คือ ค่าน้ำหนักของคำสำคัญในเอกสารเปรียบเทียบ

n คือ จำนวนคำสำคัญที่ทำการเปรียบเทียบ

1.3) วิธีการวัดความคล้ายคลึงของเอกสารแบบ Cosine เป็นวิธีการเปรียบเทียบความคล้ายคลึงของเอกสาร โดยแต่ละเอกสารจะถูกแทนด้วยเวกเตอร์ขนาดเอ็น (N-Dimensional Vector) ซึ่งเก็บค่าน้ำหนักคำแต่ละคำในเอกสารนั้น (N-Dimensional Vector in Term Space) การเปรียบเทียบความคล้ายคลึงของเอกสารจะเปรียบเทียบ โดยดูจากมุมโคไซน์ของมุมระหว่าง 2 เวกเตอร์ของเอกสาร หากเอกสารทั้งสองเอกสารคล้ายคลึงกันมากเวกเตอร์ของเอกสารทั้ง 2 จะทับกันเกือบสนิท มุมจึงมีค่าน้อย ค่าโคไซน์ที่ได้จะมีค่ามาก ค่าความคล้ายคลึงสูงสุดที่วัดด้วยวิธีนี้จะมีค่าเท่ากับ 1 โดยมีความหมายคือเวกเตอร์ทั้งสองทำมุมระหว่างกัน 0 องศา หรือเวกเตอร์ทั้งสองมีทิศทางเดียวกัน นั่นคือเอกสารทั้งสองเหมือนกัน

$$\frac{[\sum_{k=1}^{n} w_{ik} . w_{jk}]}{\sqrt{\sum_{k=1}^{n} {w_{ik}}^2} . \sqrt{\sum_{k=1}^{n} {w_{jk}}^2}}$$

เมื่อ *Wi* คือ ค่าน้ำหนักของคำสำคัญจากชุดคำค้น

Wj คือ ค่าน้ำหนักของคำสำคัญในเอกสารเปรียบเทียบ

n คือ จำนวนคำสำคัญที่ทำการเปรียบเทียบ

1.4) วิธีการวัดความคล้ายคลึงของเอกสารแบบ Jaccard เป็นวิธีการวัดค่าความคล้ายคลึงของเอกสาร ที่ถูกแทนด้วยเวกเตอร์น้ำหนักของคำที่ปรากฏในเอกสาร แล้วนำค่าความคล้ายคลึงของเวกเตอร์สองเอกสาร หารด้วยผลบวกของผลรวมค่าของเวกเตอร์สองเอกสารและลบด้วยผลรวมของผลคูณของทั้งสองเวกเตอร์

$$\frac{[\sum_{k=1}^{n} w_{ik} . w_{jk}]}{\sum_{k=1}^{n} w_{ik} + \sum_{k=1}^{n} w_{jk} - \sum_{k=1}^{n} w_{ik} . w_{jk}}$$

เมื่อ *Wi* คือ ค่าน้ำหนักของคำสำคัญจากชุดคำค้น

Wj คือ ค่าน้ำหนักของคำสำคัญในเอกสารเปรียบเทียบ

n คือ จำนวนคำสำคัญที่ทำการเปรียบเทียบ

โมเดลปริภูมิเวกเตอร์เป็นการจัดรูปแบบของเอกสารหรือตัวแทนของเอกสาร (Document Representation) เพื่อให้ง่ายต่อการประมวลผลและตีความ ซึ่งรูปแบบจะมีลักษณะของการแสดงความสัมพันธ์ระหว่างคำในเอกสารทั้งหมดด้วยเวกเตอร์ 2 มิติ ซึ่งคำที่ได้นั้น

ต้องผ่านกระบวนการ ทำให้เป็นบรรทัดฐาน (Normalization) แล้ว เช่น การตัดคำหยุดออกไป รวมถึงการให้น้ำหนักกับคำเหล่านั้น รูปแบบเช่นนี้เรียกว่าการจัดรูปแบบโมเดลปริภูมิเวกเตอร์หรือรูปแบบของเอกสารแบบถุงของคำ (Bag of Words) ดังตารางที่ 2.6

ตารางที่ 2.6 การจัดรูปแบบของเอกสารแบบถุงของคำ

	W1	**W2**	**…**	**Wk**	**…**	**Wv**
D1	W11	W12	…	W1k	…	W1v
D2	W21	W22	…	W2k	…	W2v
…	…	…	…	…	…	…
Dn	Wn1	Wn2	…	Wnk	…	Wnv

จากตารางที่ 2.6 เป็นลักษณะการอธิบายโมเดลปริภูมิเวกเตอร์เป็นทำการจัดเก็บน้ำหนักคำค้น (Wi) ของแต่ละเอกสาร (Di) โดยแสดงน้ำหนักของคำสำคัญที่อยู่ในเอกสารแต่ละรายการเป็นชุดเชื่อมต่อกันเป็นเส้นตรง (Vector) พร้อมกับกำหนดค่านำหนักให้กับคำสำคัญแต่ละคำตามความต้องการแล้ว จะทำการเปรียบเทียบชุดคำศัพท์ที่เหมือนหรือสอดคล้องกันเหล่านั้นจนครบทุกคน เพื่อนำมาคำนวณเป็นผลรวมของความคล้ายคลึงของแต่ละเอกสาร

2) รูปแบบการทำงานของโมเดลปริภูมิเวกเตอร์

2.1) ให้ความสำคัญกับความถี่ของคำที่ปรากฏอยู่ในเอกสารและความถี่มีผลต่อการให้น้ำหนักของคำในเอกสาร โดยที่

2.1.1) การนับความถี่ของคำ (Term Frequency: TF) คือ การใช้ความถี่ของคำ เช่น พบ 1 ครั้งเรียกว่า เทอม (Term) ทั้งนี้ขึ้นอยู่กับจำนวนคำของเอกสาร โดยเทอมจะแทนคำศัพท์ของแต่ละคำ

2.1.2) ค่าน้ำหนักของคำ (Term Weight: W) ความถี่ของคำคำหนึ่งที่พบในทุก ๆ เอกสาร

2.1.3) การสร้างตารางความถี่ของคำ เป็นขั้นตอนในการหาความถี่ของคำศัพท์ และกำหนด ค่าน้ำหนักของคำศัพท์ในแต่ละเอกสารลงในตารางความถี่ของคำ

2.1.4) การนำเสนอข้อมูลในเชิงเวกเตอร์ ซึ่งคำศัพท์ในตารางความถี่ของคำจะถูกนำเสนอ ในเชิงเวกเตอร์ โดยจะถูกมองเป็นแถวลำดับของเวกเตอร์ เช่น (1, 1, 0, 0, 0)

2.1.5) การคำนวณค่าความคล้ายคลึง โดยคุณสมบัติของเวกเตอร์ทำให้สามารถคำนวณค่าความคล้ายคลึงระหว่างเวกเตอร์ได้จากค่าสัมประสิทธิ์ต่าง ๆ เช่น ค่าสัมประสิทธิ์โคไซน์ของมุมระหว่างคู่เวกเตอร์ ซึ่งจะมีค่าอยู่ระหว่าง 0 ถึง 1

ข้อดีของโมเดลปริภูมิเวกเตอร์ คือใช้คณิตศาสตร์เรียบง่ายในการคิด มีการพิจารณาจากความถี่ของคำและสามารถจัดลำดับของเอกสารได้และสามารถใช้กับเอกสารที่มีข้อมูลมาก ๆ ได้ดี ซึ่งโมเดลเวกเตอร์เป็นการจับคู่แบบไม่เที่ยงตรงหรือสามารถแยกรายการคำค้นเป็นเทอมได้ ทำให้สามารถค้นหาแบบเป็นบางส่วน (Partial Search) ได้

ข้อเสียของโมเดลปริภูมิเวกเตอร์ คือ โมเดลปริภูมิเวกเตอร์จะไม่สนใจความหมายของคำ วลี โครงสร้างของคำ หรือคำที่มีความหมายเหมือนกัน เพราะโมเดลปริภูมิเวกเตอร์อยู่บนข้อสมมุติที่ว่า คำทุกคำไม่มีความสัมพันธ์กัน ซึ่งในความเป็นจริงคำจะมีความสัมพันธ์หรืออ้างอิงกันในเชิงความหมายได้

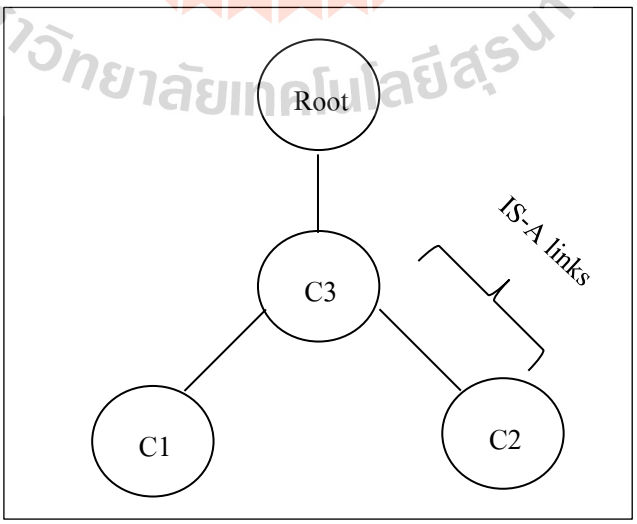
วิธีการข้างต้นเป็นต้นแบบทฤษฎีที่มักใช้ในการหาความคล้ายคลึงของเอกสารระหว่างเอกสาร หรือระหว่างคำค้นกับเอกสารในระบบ ในงานวิจัยนี้ก็ได้นำแนวคิดในการเปรียบเทียบคำค้นและ คำสำคัญในเอกสารเพื่อหาค่าความคล้ายคลึงเช่นกัน และได้เพิ่มแนวคิดการวัดความคล้ายคลึงเชิงความหมายเพิ่มเข้าไปในกระบวนการค้นคืนด้วย เพื่อเพิ่มประสิทธิภาพของการเปรียบเทียบความคล้ายคลึงตามโมเดลปริภูมิเวกเตอร์ โดยจะนำเสนอในลำดับถัดไป

2.4.2.2 การวัดความคล้ายคลึงเชิงความหมาย (Semantic Similarity Measurement)

วิธีวัดความคล้ายของคำที่มีความหมายคล้ายกัน ในที่นี้จะแบ่งออกเป็น 2 กลุ่มคือ

1) การวัดระยะโหนด (Edge Counting Methods)

การวัดระยะโหนด คือ การวัดระยะคำศัพท์ 2 คำ โดยใช้ความยาวของ path เช่น ความสัมพันธ์แบบ IS-A links ที่เชื่อมต่อแต่ละคำ และตำแหน่งของคำในกลุ่ม ซึ่งงานวิจัยของกลุ่มนี้ได้แก่ Claudia and Chodorow (1998) ได้ทำการวัดความคล้ายคลึงกันของคอนเซ็ปต์ C1 และ C2 โดยใช้ระยะห่างระหว่าง path (IS-A links) C1 และ C2 ซึ่งคอนเซ็ปต์ C1 และ C2 ต้องเป็นคำนามเท่านั้น โดยการวิเคราะห์โครงสร้างของคลังคำศัพท์ที่สร้างความสัมพันธ์กัน ดังรูปที่ 2.13 แสดงลักษณะของความสัมพันธ์ของคลังคำ



รูปที่ 2.14 โครงสร้างความสัมพันธ์ของคลังคำตัวอย่าง

จากรูปที่ 2.14 แสดงถึงตัวอย่างโครงสร้างของคลังคำซึ่งวิธีของ Claudia and Chodorow (1998) จะคำนวณโดยใช้ระยะห่างระหว่าง path (IS-A links) C1 และ C2 ซึ่งมีสมการคำนวณคือ

$$sim_{LC}(c_1 c_2) = -log \frac{len(c_1 c_2)}{2L} \qquad (2\text{-}1)$$

เมื่อ c1 และ c1 คือ คำใด ๆ ที่ต้องการหาค่าความคล้ายกัน

len(c1 ,c2) คือ จำนวนของ IS-A links ระหว่าง c1 ถึง c2

L คือ ความลึกโดยรวมของชุดคลังคำ

Wu and Palmer (1994) ได้ทำการวัดความคล้ายของคอนเซ็ปต์ C1 และ C2 โดยคำนวณ link ที่ขึ้นตรงต่อกัน (IS-A) ระหว่างจากคอนเซ็ปต์ที่เป็นซุปเปอร์คลาสของ C1 และ C2 โดยจากรูปที่ 2.14 จะคำนวณด้วยสมการ

$$sim_{WuP}(c_1 c_2) = \frac{2.N_3}{N_1 + N_2 + 2\ .N_3} \qquad (2\text{-}2)$$

เมื่อ c1 และ c1 คือ คำใด ๆ ที่ต้องการหาค่าความคล้ายกัน

N1 คือ จำนวนของ IS-A links จาก c1 ถึงคำที่เชื่อมกันใกล้สุดระหว่าง c1 และ c2

N2 คือ จำนวนของ IS-A links จาก c2 ถึงคำที่เชื่อมกันใกล้สุดระหว่าง c1 และ c2

N3 คือ จำนวนของ IS-A links ระหว่าง คำที่เชื่อมกันใกล้สุระหว่าง c1 และ c2 ถึง root ในโครงสร้างของอนุกรมวิธาน

2) การวัดจากเนื้อหาสารสนเทศ (Information Content Methods)

วัดความแตกต่างของเนื้อหาของคำค้นทั้งสองคำ โดยใช้ความเป็นไปได้ที่จะเกิดขึ้นในเอกสาร ซึ่งงานวิจัยของกลุ่มประเภทนี้ได้แก่ Resnik (1995) ได้พิจารณาคอนเซ็ปต์แต่ละคอนเซ็ปต์ว่ามี ค่าเชิงเนื้อหา (Information Content) ไม่เท่ากัน ซึ่งค่าเชิงเนื้อหาคิดได้จาก ค่าทางสถิติความน่าจะเป็นในการปรากฏของคอนเซ็ปต์นั้น (รวมถึงคอนเซ็ปต์ที่เป็นตัวสืบทอดจากคอน

มหาวิทยาลัยเทคโนโลยีสุรนารี

เซ็ปต์นั้นด้วย เช่น ซับคลาส) เปรียบเทียบกับค่าสถิติของการเกิดคอนเซ็ปต์ทุกคอนเซ็ปต์ภายในชุดข้อมูลที่เตรียมไว้ (Dataset) โดยมีสูตรการคำนวณดังนี้

$$IC(concopt) = -\log(P(concept)) \qquad (2\text{-}3)$$

ซึ่งค่า P(concept) คือค่าความน่าจะเป็นที่คอนเซ็ปต์นั้นจะปรากฏในคลังข้อมูล (Corpus) หาได้จากสูตรดังต่อไปนี้

$$P(concept) = freq(concept)/freq(root) \qquad (2\text{-}4)$$

และถ้าวัดความคล้ายคลึงกันระหว่าง C1 และ C2 โดยมี C0 เป็นคำที่เชื่อมกันใกล้สุดระหว่าง C1 และ C2 (Lowest Common Subsume) แล้วจะสามารถคำนวณด้วยสูตร

$$sim_R(c_1, c_2) = \frac{2.\log p(c_0)}{\log p(c_1)+\log p(c_0)} \qquad (2\text{-}5)$$

หรือสูตรคำนวณดังนี้

$$sim_R(c_1, c_2) = -\log(\, lcs(c_1, c_2)) \qquad (2\text{-}6)$$

โดย lcs (c1,c2) คือ คำที่เชื่อมกันใกล้สุดระหว่าง C1 และ C2 (lowest common subsume)

Jiang and Conrath (1997) ได้ใช้หลักการการวัดความคล้ายคลึงกันเหมือนของ Resnik แต่วิธีวัดระยะห่างเชิงความหมายจะผกผันกับค่าความคล้ายกัน

$$dist(concept_1, concept_2) = IC(concept_1 + concept_2) - 2 * (IC(LCS) \qquad (2\text{-}7)$$

$$sim_{JnC} = \frac{1}{IC(concept_1)+IC(concept_2)-2*IC(LCS)} \qquad (2\text{-}8)$$

โดย IC คือ ค่าเชิงข้อมูลซึ่งสามารถคำนวณได้จากสมการ (2-3)

LCS คือ คำที่เชื่อมกันใกล้สุดระหว่าง concept1 และ concept2 (lowest common subsume)

มหาวิทยาลัยเทคโนโลยีสุรนารี

Jiang and Conrath (1997) ได้ใช้ วิธีหาค่า Information Content เพื่อนำมาหาค่า ระยะห่างของ path ระหว่างคอนเซ็ปต์ ที่ต้องการหาความคล้ายคลึงกัน ซึ่งคำที่ใช้ต้องเป็นคำนามและผลนั้นจะมีค่าตั้งแต่ 0 ขึ้นไป

Lin (1998) ใช้แนวคิดสมการคล้ายของ Wu and Palmer (1994) แต่ใช้วิธีคิดแบบ ใช้ค่าเชิงข้อมูลนำมาคำนวณ

$$sim_{Lin} = \frac{2.IC(LCS)}{IC(concept_1)+IC(concept_2)} \qquad (2\text{-}9)$$

โดย IC คือ ค่าเชิงข้อมูลซึ่งสามารถคำนวณได้จากสมการ(3)

LCS คือ คำที่เชื่อมกันใกล้สุดระหว่าง concept1 และ concept2

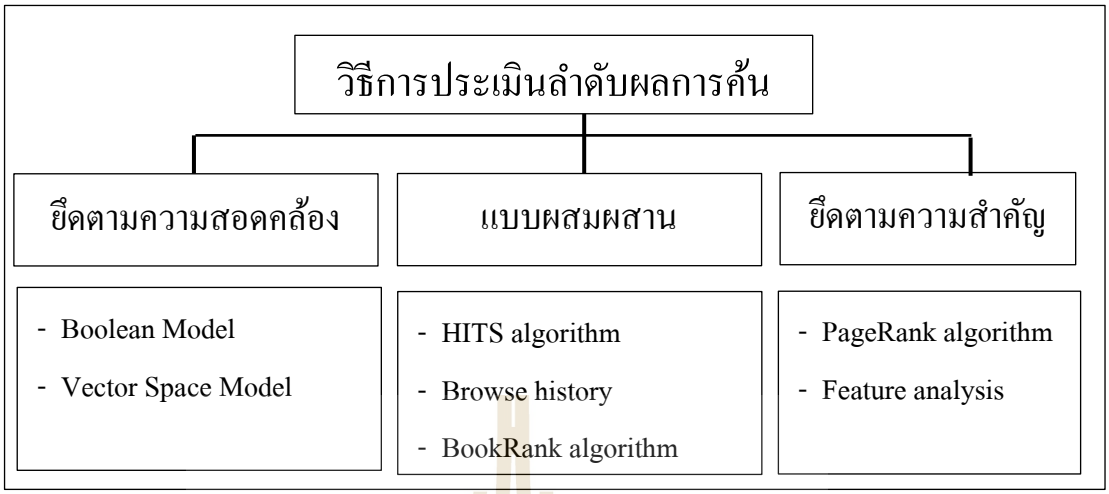
วิธีการวัดความคล้ายของ Lin ผลลัพธ์ที่ได้จะมีค่าอยู่ระหว่าง 0 - 1 เท่านั้น

จากแนวคิดการวัดความคล้ายคลึงเชิงความหมายข้างต้น ผู้วิจัยได้ประมวลแนวคิดและสูตรการคำนวณโดยวิเคราะห์ถึงองค์ประกอบโครงสร้างข้อมูลและการใช้คลังคำเพื่อช่วยเพิ่มการค้นคืน เชิงความหมายในการวัดความคล้ายคลึงของเอกสารและคำค้นจากผู้ใช้ จึงเลือกใช้แนวคิดการวัดความคล้ายคลึงเชิงความหมายของลิน (Lin, 1998) เพื่อประเมินความคล้ายคลึงของคำในเชิงความหมาย หรือระยะห่างของคำในคลังคำ อีกทั้งประยุกต์แนวคิดจากการวัดระยะห่างของคอนเซ็ปต์ของ Wu and Palmer (1994) ร่วมด้วยเพื่อนำสูตรทั้งสองมาประยุกต์ในการวัดความคล้ายคลึงทั้งด้านระยะทางของคอนเซ็ปต์ที่ค้นหาเพิ่มเติม นอกจากนี้ ผู้วิจัยยังใช้คลังคำศัพท์ช่วยการค้นคืนในเชิงความหมายพร้อมกับการใช้ค่าน้ำหนักของคำที่ปรากฏในเมทาดาทาของเอกสารที่มีค่าน้ำหนักที่แตกต่างกันออกไป ตามการค้นคืนของผู้ใช้และความสำคัญของตำแหน่งที่พบคำนั้นเพิ่มเป็นค่าคะแนนให้กับผลการค้นคืนด้วย

2.4.3 การจัดลำดับผลการค้นคืน (Search Result Ranking)

2.4.3.1 เทคนิคในการประเมินคุณภาพของผลการค้นคืน

เทคนิคในการประเมินคุณภาพของผลการค้นคืน เพื่อการจัดเรียงลำดับรายการผลลัพธ์ที่ได้จากการค้นคืนของระบบค้นคืนสารสนเทศนั้น สามารถสรุปได้เป็น 3 แนวทางหลัก ตามรูปที่ 2.15 ดังนี้



รูปที่ 2.15 เทคนิคการจัดเรียงลำดับผลการค้นคืน

1) จัดเรียงลำดับตามความสอดคล้อง (Relevance) วิธีการนี้ระบบจะรับคำถามจากผู้ใช้งาน (Query dependent) จากนั้นจะนำไปคำนวณค่าความสัมพันธ์ระหว่างคำถามกับคำสำคัญในทุก เขตข้อมูล หรือองค์ประกอบย่อยเมทาดาทา โดยอาจพิจารณาความสอดคล้องกันจากปริมาณความถี่ที่มีคำสำคัญจากคำถามปรากฏอยู่ในเอกสารหรือตัวแทนเอกสาร แล้วจึงจัดเรียงลำดับตามความสอดคล้องจากมากไปน้อย ซึ่งเทคนิคดังกล่าวนี้เป็นเทคนิคที่ใช้ในระบบการค้นคืนสารสนเทศแบบเดิม ยกตัวอย่าง เช่น แบบจำลองบูลีน (Boolean model) และแบบจำลองปริภูมิเวกเตอร์ (Vector Space model) (Baeza-Yates and Ribeiro-Neto, 2005)

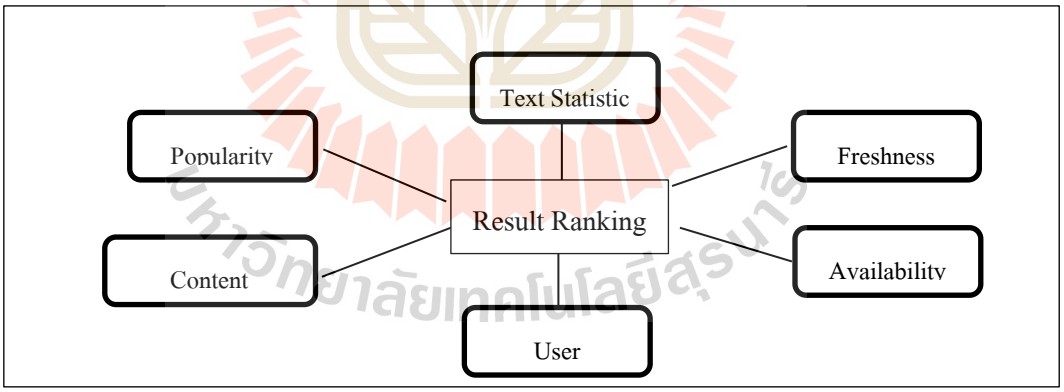
2) จัดเรียงลำดับตามความสำคัญ (Importance) วิธีการนี้จะวิเคราะห์ส่วนประกอบต่าง ๆ ของเอกสารเพื่อประมาณค่าความสำคัญของข้อมูลเหล่านั้นไว้ก่อนล่วงหน้าโดยไม่พิจารณาถึงคำถามของผู้ใช้งาน (Query independent) เป็นการวิเคราะห์โครงสร้างการเชื่อมโยงกันของเนื้อหาและเมทาดาทาในชุดข้อมูล เช่น อัลกอริทึมเพจแร็งค์ (Brin, and Page, 1998; Page and et al., 1999) หรือวิเคราะห์ลักษณะเด่นขององค์ประกอบต่าง ๆ ที่ปรากฏบนหน้าเอกสาร เช่น อัลกอริทึมแอลพีอี (Linear Page Estimation: LPE) อัลกอริทึมดอม (Document Object Model: DOM) (Chen and et al., 2001) และอัลกอริทึมวิปส์ (Vision-based Page Segmentation: VIPS) (Cai, and et al., 2003; Song and et al., 2004) เป็นต้น หลังจากนั้นรายการผลการค้นจึงถูกจัดเรียงลำดับด้วยค่าความสำคัญเหล่านี้

3) จัดเรียงลำดับตามความสำคัญโดยพิจารณาเฉพาะรายการและเขตข้อมูลที่สนใจหรือสอดคล้องกับคำถามของผู้ใช้ กล่าวคือ วิธีการนี้จะพิจารณาผลลัพธ์ที่ได้จากการค้นคืน

ตามค่าความสอดคล้องกันคำถามของผู้ใช้งาน ขยายขอบเขตของข้อมูลโดยพิจารณาเมทาดาทาที่ถูกเชื่อมโยง (linked Metadata) หลังจากนั้นจึงวิเคราะห์โครงสร้างการเชื่อมโยงกันเพื่อคำนวณค่าความสำคัญ เช่น อัลกอริทึมฮิตส์ (Hyperlink-Induced Topic Search: HITS) (Kleinberg, 1999) นอกจากนี้ ยังมีเทคนิคที่อาจจัดอยู่ในแนวทางการผสมผสานระหว่างการจัดเรียงที่ขึ้นต่อ และไม่ขึ้นต่อคำถามของผู้ใช้งาน (Hybrid) นั่นคือ การทำเหมืองข้อมูลด้านประวัติการเรียกชมหน้าเว็บเพจของผู้ใช้งาน เช่น อัลกอริทึมบราวซ์แร็งค์ (BrowseRank algorithm) (Lui et al., 2008, 2010) แทรฟฟิคเวทแร็งค์ (Traffic-weighted Rank algorithm) (Meiss et al., 2008) และบุคแร็งค์ (BookRank algorithm) (Gonçalves et al., 2009) เป็นต้น ถึงแม้ว่าเทคนิคดังกล่าวนี้จะทำให้สามารถวิเคราะห์ค่าความสำคัญให้กับเว็บเพจไว้ก่อนล่วงหน้าได้ก็ตาม หรือที่เรียกว่า “รูปแบบออฟไลน์” อย่างไรก็ดี การทำเหมืองข้อมูลจำเป็นจะต้องทราบความสัมพันธ์ระหว่างคำถามและลำดับเว็บเพจผลลัพธ์ที่ถูกเรียกดูโดยผู้ใช้งานเป็นจำนวนที่มากพอ จึงจะทำให้การวิเคราะห์มีความน่าเชื่อถือ

2.4.3.2 ปัจจัยที่มีผลต่อการจัดอันดับผลการค้นคืน

นักวิจัยได้จำแนกหมวดหมู่ของปัจจัยที่มีผลต่อการจัดอันดับของระเบียนรายการผลการค้นคืนเป็น 6 หมวดหมู่ ซึ่งในแต่ละหมวดหมู่จะมีปัจจัยย่อยภายในแต่ละหมวด (Weideman, 2009; Lewandowski, 2009) ตามที่แสดงในรูปที่ 2.16



รูปที่ 2.16 ปัจจัยที่มีผลต่อการจัดอันดับของระเบียนผลการค้นคืน (Lewandowski, 2009)

1) การเก็บสถิติของข้อความในเนื้อหา (Text Statistics) เป็นปัจจัยหลักที่ใช้พิจารณาความเกี่ยวข้องของผลการค้นในระบบการค้นคืนสารสนเทศแบบดั้งเดิม โดยจะเน้นที่คำสำคัญหรือข้อความที่ใช้ค้นเป็นหลัก หรือคำดัชนี ถ้าคำค้นที่ใช้ตรงกับคำดัชนีที่เป็นตัวแทนของเอกสาร หรือเนื้อหาในตัวเอกสาร เอกสารจะถูกเรียกขึ้นมาแสดง กระบวนการนี้เรียกว่าการ

เปรียบเทียบข้อความ (Text Matching) อย่างไรก็ดี ปัจจัยที่ใช้จัดอันดับผลการค้นคืนเพียงอย่างเดียวนั้นไม่ได้บ่งชี้ถึงคุณภาพของตัวเอกสารโดยตรง แต่จะช่วยเพิ่มความน่าเชื่อถือของผลการค้นคืนโดยรวมของระบบ ปัจจัยย่อยในการเปรียบเทียบความคล้ายคลึงในกลุ่มนี้สามารถแบ่งออกด้วยวิธีหลายประการ ได้แก่

1.1) ความถี่และค่าน้ำหนักของคำ (Term Frequency and Term Weighting) ได้แก่ ค่าความถี่ของคำดัชนีใดที่ปรากฏในเอกสารและค่าความถี่ของคำดัชนีใด ๆ ที่ปรากฏในเอกสารทั้งหมดในระบบ (Term Frequency/Invert Document Frequency: TF/IDF) โดยค่านี้จะช่วยประเมินได้ว่าเอกสารใดที่มีคำสำคัญและมีจำนวนมากพอที่จะประเมินได้ว่าในเอกสารนั้นมีเนื้อหาเกี่ยวกับเรื่องนั้นจริง ไม่ใช่เป็นเพียงคำประกอบหรือเรื่องย่อย (Larson, 2012: p. 21; Baeza-Yates and Ribeiro-Neto, 2011) ด้วยการนับความถี่นี้จะช่วยในการค้นคืนค้นแบบบางส่วนของคำค้น (Partial Search) ซึ่งดีกว่าการใช้งานตรรกะแบบบูลีนที่ต้องเปรียบเทียบทุกคนค้น (Khoo and Wan, 2004: p. 112)

1.2) ลำดับคำค้น ระยะห่างและตำแหน่งของคำในเอกสาร (Search Term Order Distance and Position in Document) เป็นอีกแนวทางหนึ่งในการคำนวณค่าเพื่อใช้จัดลำดับรายการผลการค้นคืน กรณีนี้จะใช้เมื่อมีการใช้คำค้นหลายคำในการค้นคืนข้อมูล การเปรียบเทียบความคล้ายคลึงของคำแต่ละคำกับรายละเอียดในเอกสารเพื่อบ่งบอกว่าเอกสารใดที่มีคำสำคัญใกล้เคียงกับคำค้นมากกว่า อาจจะเป็นเอกสารที่ผู้ใช้ต้องการมากกว่าด้วย นอกจากนี้ ตำแหน่งที่คำสำคัญปรากฏอยู่ในเอกสารจะช่วยในการให้คะแนนความสำคัญหรือความเกี่ยวข้องกับเนื้อหาได้เช่นกัน เช่น ปรากฏในชื่อเรื่อง หรือปรากฏในบทคัดย่อย่อมมีคะแนนที่สูงกว่าการปรากฏในเนื้อหาทั่วไป เป็นต้น (Dopichaj, 2009)

2) การได้รับความนิยม (Popularity) การใช้ข้อมูลที่บ่งบอกถึงความนิยมในรายการทรัพยากรสารสนเทศเป็นอีกแนวทางหนึ่งที่ช่วยเพิ่มความมั่นใจให้ผู้ใช้เมื่อพบรายการสารสนเทศดังกล่าว เนื่องจากความนิยมที่เก็บข้อมูลมานั้นจะเป็นข้อมูลจากผู้ใช้คนอื่น ๆ มากกว่าการประเมินที่ตัวเนื้อหาหรือข้อมูลในทรัพยากรสารสนเทศ (Surowiecki, 2005) ปัจจัย หรือตัวบางชี้ในประเด็นนี้ ได้แก่

2.1) ความถี่ในการเข้าใช้ (Usage Frequency) การนับความถี่ของการใช้งานเป็นข้อมูลที่นิยมใช้อีกข้อมูลหนึ่ง เนื่องจากสามารถแสดงถึงความน่าสนใจของทรัพยากรสารสนเทศได้ โดยการนับความถี่การเข้าใช้เป็นการนับที่การเปิดลิงค์หรือเปิดรายการเอกสารนั้นผ่านทางระบบ หรือนับจากจำนวนการดาวน์โหลดไฟล์ต้นฉบับไปใช้งานในกรณีที่เป็นสื่อดิจิทัล ซึ่งการนับการใช้งานอาจจะมีความคลาดเคลื่อนในกรณีที่จำนวนความถี่มีสูงมาก ทำให้ผู้ใช้เกิดความเข้าใจผิดและ

เลือกเปิดเอกสารนั้น ๆ ตามจำนวนที่แสดงความถี่ ซึ่งเอกสารหรือข้อมูลนั้นอาจจะไม่ได้มีคุณภาพเสมอไป แต่มีการนับความถี่การเปิดเพิ่มขึ้นเรื่อย ๆ ไปแล้วก็ได้

2.2) ค่าความนิยม (Rating) การให้คะแนนรายการทรัพยากรสารสนเทศ หรือเรียกว่าการบอกเล่าประสบการณ์การใช้ หรือความรู้สึกต่อสิ่งหนึ่งสิ่งใดที่มีต่อบุคคลคนหนึ่ง ๆ เป็นสิ่งที่ผู้ใช้ปัจจุบันต้องการและผู้ให้บริการพยายามสกัดค่าความนิยมนี้จากผู้ใช้อยู่เสมอ ๆ ซึ่งการให้ค่าคะแนนความนิยมจะมีรายละเอียดเพิ่มมากกว่าการนับความถี่การเข้าใช้ ผู้ใช้สามารถแสดงความคิดเห็นหรือสะท้อนคุณภาพของเนื้อหาได้ดีกว่า ถึงแม้จำนวนการใช้จะมาก แต่ถ้าค่าความนิยมต่ำก็แสดงถึงรายการทรัพยากรนั้นไม่เป็นไปตามที่ผู้ใช้ต้องการ

2.3) การอ้างถึงและการถูกอ้างถึง (Reference Counting) การอ้างอิงเป็นส่วนช่วยเพิ่มความน่าเชื่อถือและคุณภาพของชิ้นงานได้เป็นอย่างดี ทรัพยากรสารสนเทศประเภทบทความวิชาการ งานวิจัย การอ้างอิงเป็นสิ่งที่ช่วยให้ผู้อ่านสามารถติดตามองค์ความรู้เพิ่มเติมได้ และเป็นมารยาทในการสร้างผลงานวิชาการที่ต้องมีเสมอ ทรัพยากรสารสนเทศที่มีจำนวนการถูกอ้างถึงมากย่อมแสดงถึงคุณภาพ ข้อมูลที่ถูกต้อง การนำเสนอที่มีประโยชน์ต่อผู้ใช้งานได้ ดังนั้น ถ้าบทความใดถูกอ้างถึงจำนวนมาก แสดงถึงบทความนั้น “น่าจะ” มีเนื้อหาและข้อมูลที่เป็นที่ยอมรับมากกว่า นอกจากบทความนั้นจะถูกอ้างถึงแล้ว ภายในบทความเองถ้ามีการอ้างอิงทรัพยากรอื่นที่มีความน่าเชื่อถือด้วย ก็แสดงถึงความตั้งใจของผู้เขียนอย่างมากที่จะค้นหาและนำเสนอข้อมูลที่มีความทันสมัยและสอดคล้องกับเรื่องที่นำเสนอ

2.4) สิทธิ์ความเป็นเจ้าของ (Owner Authority) เป็นความน่าเชื่อถือที่ติดมากับตัวทรัพยากรสารสนเทศ และเป็นเหมือนบรรจุภัณฑ์ที่มีป้ายรับรองคุณภาพปิดไว้ ข้อมูลเหล่านี้ ได้แก่ ชื่อเจ้าของผลงาน ผู้เผยแพร่ และสำนักพิมพ์ มีหลายครั้งที่ผู้ใช้ติดตามผลงานโดยดูจากชื่อผู้เขียน หรือสำนักพิมพ์เพราะเชื่อมั่นในคุณภาพและความสมบูรณ์ของเนื้อหาที่นำเสนอ ดังนั้น รายละเอียดเกี่ยวกับผู้เป็นเจ้าของผลงานจึงมีส่วนในการจัดอันดับการแสดงผลได้เช่นกัน

3) ความทันสมัย (Freshness) แสดงถึงความใหม่ของเนื้อหาและข้อมูลภายในทรัพยากรสารสนเทศที่มีความเป็นปัจจุบัน หรือทันเหตุการณ์ เช่น บทความด้านเทคโนโลยี สถิติ หรือนวัตกรรม ความใหม่ของเนื้อหาบทความมีผลอย่างมากกับการตัดสินใจเลือกใช้งานของผู้ใช้ ความทันสมัยสามารถดูได้จากปัจจัย 2 ส่วนคือ ปีพิมพ์หรือวันที่เผยแพร่ (Publication Date) แสดงถึงวันที่เริ่มนำเสนอผลงานนั้น ๆ และวันที่สามารถเข้าถึงได้ (Accession Date) หมายถึงวันที่ข้อมูลเหล่านั้นมีผู้เข้าใช้งานหรือมีการเข้าใช้เป็นปกติ

4) ความพร้อมใช้ (Availability) โดยมากแล้วระบบสารสนเทศจะเก็บข้อมูลไว้ในพื้นที่ที่จัดเตรียมไว้เพื่อให้ผู้ใช้สามารถเข้าถึงได้ทันทีที่ต้องการ แต่ปัจจุบันได้มีแนวคิดการใช้งานทรัพยากรร่วมกัน ทำให้ทรัพยากรบางประเภทที่มีจำนวนน้อย ราคาสูง ต้องนำมาแบ่งปันกันใช้งาน ทำให้ต้องแปลงทรัพยากรสารสนเทศเหล่านั้นเป็นรูปแบบดิจิทัลควบคู่กับทรัพยากรที่เป็นแบบกายภาพ หรือเป็นตัวเล่มจริงไปด้วย แต่ทั้งนี้ การเข้าถึงไฟล์ดิจิทัลมีข้อจำกัดในเรื่องของระบบเครือข่าย การเก็บไฟล์ไว้นอกระบบ หรือการแบ่งปันใช้งานร่วมกัน ซึ่งอาจจะมีผลกับการเข้าถึงไฟล์ที่ไม่พร้อมใช้ได้ ส่วนเอกสารตัวเล่ม (Physical Object) สามารถใช้งานได้ทันทีไม่ต้องใช้เทคโนโลยีเครือข่ายร่วมด้วย แต่ด้วยจำนวนที่มีจำกัด หรืออยู่คนละพื้นที่ก็ทำให้ทรัพยากรเหล่านั้นไม่ดึงดูดให้ผู้ใช้สนใจมาใช้งาน ในการจัดอันดับนำเสนอสามารถเรียบเรียงรายการที่พร้อมใช้ก่อนเพื่อให้ผู้ใช้ไม่เสียเวลาในการเลือกรายการที่ไม่พร้อมให้บริการ หรือไม่สามารถเข้าถึงได้ในเวลานั้น

5) รายละเอียดของเนื้อหา (Content Properties) รายละเอียดของเนื้อหาเป็นส่วนประกอบหลักที่ผู้ใช้พิจารณาเมื่อรายการผลการค้นคืนแสดงบนหน้าจอ รายละเอียดของเนื้อหาที่แสดงนั้นเป็นข้อมูลตัวแทนสารสนเทศที่จัดทำตามเมทาดาทาที่กำหนดไว้ ดังนั้น ผู้พัฒนาระบบหรือผู้ออกแบบโครงสร้างเมทาดาทาสามารถเลือกรายการข้อมูลที่จำเป็นสำหรับการตัดสินใจของผู้ใช้เพื่อสนับสนุนการตัดสินใจและพิจารณาเลือกใช้ทรัพยากรแต่ละรายการได้ ข้อมูลที่ผู้ใช้มักนำมาใช้ในการพิจารณาได้แก่ ภาษาของเอกสาร (Language) รูปแบบของไฟล์ดิจิทัล (Format) เป็นต้น

6) ภูมิหลังของผู้ใช้ (User Background) ผู้ใช้งานระบบสารสนเทศมีหลายกลุ่ม ซึ่งแต่ละกลุ่มมีระดับความสามารถในการใช้งานระบบสารสนเทศเพื่อค้นคืนข้อมูลไม่เท่ากัน โดยแต่ละกลุ่มจะมีปัจจัยที่แฝงอยู่ภายในหลายประการณ์ เช่น ความคุ้นเคย (Familiar) ประสบการณ์การใช้งาน (User Experience) หรือภูมิหลังของผู้ใช้ เช่น ความชอบ ความสนใจ และขอบเขตงานที่รับผิดชอบ ซึ่งได้กลายเป็นสิ่งที่ทำให้รูปแบบพฤติกรรมการใช้งานระบบแตกต่างกัน ผู้พัฒนาระบบสามารถนำลักษณะของผู้ใช้งานไปเป็นปัจจัยหนึ่งในการคัดกรองผลการค้นคืนให้เหมาะสมกับลักษณะเฉพาะของผู้ใช้ หรือพัฒนาระบบให้สามารถรองรับการปรับเปลี่ยนการแสดงผลได้หลากหลายวิธี

จากการประมวลแนวคิดในการจัดอันดับผลการค้นคืน จะเห็นว่าการจัดอันดับสามารถนำผลคะแนนจากการค้นคืน เช่น คะแนนความคล้ายคลึงเชิงความหมาย และข้อมูลตามโครงสร้างเมทาดาทาบางรายการมาเป็นปัจจัยในการจัดอันดับได้ เช่น ประเภทของสื่อ ภาษาของเอกสาร หมวดหมู่ หัวเรื่อง เป็นต้น ดังนั้น ในงานวิจัยชิ้นนี้จึงใช้แนวคิดการจัดอันดับข้างต้นทั้งด้าน

สถิติของข้อความในเนื้อหา ความนิยม ความทันสมัย ความพร้อมใช้ รายละเอียดตัวแทนสารสนเทศ และภูมิหลังของผู้ใช้มาเป็นปัจจัยในการพัฒนาโมเดล โดยออกแบบให้ผู้ใช้สามารถปรับเปลี่ยนรายละเอียดการแสดงผลของเขตข้อมูลตามโครงสร้างเมทาดาทาได้ และเลือกการจัดอันดับตามสถิติการใช้งานหรือรายละเอียดของสารสนเทศตามที่ผู้ใช้ต้องการได้ด้วย

2.5 งานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ เป็นการวิจัยและพัฒนาเพื่อคิดค้นกระบวนการเพื่อช่วยให้การค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทามีความสะดวกและตรงกับความต้องการของผู้ใช้ โดยใช้ครอว์เลอร์ (Crawler) ในการรวบรวมข้อมูลเอกสารอาร์ดีเอฟตามยูอาร์ไอของข้อมูลที่เกี่ยวข้องกับเลิร์นนิงอ็อบเจกต์ และทำการวิเคราะห์เปรียบเทียบและผสานชุดเมทาดาทาให้เป็นโครงสร้างเดียวกัน จากนั้นทำการลดความซ้ำซ้อนของข้อมูลที่เชื่อมโยงไปยังยูอาร์ไอเดียวกันด้วยการวิเคราะห์รายการยูอาร์ไอแต่และรายการเพื่อจัดกลุ่มเอกสารที่เป็นเอกสารเดียวกันหรือเหมือนกันโดยการวิเคราะห์ข้อมูลภายในชุดเมทาดาทา ส่วนสุดท้ายคือ การออกแบบการค้นคืนเลิร์นนิงอ็อบเจกต์จากการใช้คำค้นของผู้ใช้ผ่านระบบการแปลงประโยคคำถามเพื่อความสะดวกในการใช้งาน

2.5.1 สรุปงานวิจัยที่เกี่ยวข้องได้ ดังนี้

โธพิคอฟและคณะ (Toupikov et al., 2009) นำเสนอวิธีการการจัดอันดับความสำคัญของรายการเอกสารภายในโอเพนดาทา เพื่อแก้ปัญหาความยุ่งยากในการตัดสินใจเลือกใช้เอกสาร เนื่องจากจำนวนข้อมูลที่อยู่ในโอเพนดาทามีปริมาณมากและเพิ่มมากขึ้นเรื่อย ๆ ทำให้ผลการค้นคืนรายการเอกสารที่ต้องการมีปริมาณมาก และแต่ละรายการก็ไม่มีข้อมูลอะไรที่บอกถึงความน่าสนใจหรือแสดงถึงความสมบูรณ์ของเนื้อหาภายใน ดังนั้น การจัดลำดับความสำคัญของผลการค้นจะช่วยให้ผู้ใช้ตัดสินใจเลือกชุดข้อมูลที่มีคุณภาพมากว่าได้ โธพิคอฟและคณะ ได้ออกแบบวิธีการจัดอันดับ (Ranking) ให้กับชุดข้อมูลด้วยการใช้เพ็จแรงก์ (Page Rank) โดยวิเคราะห์จากความสัมพันธ์ระหว่างชุดข้อมูล หากชุดข้อมูลใดมีการสร้างความสัมพันธ์ระหว่างกัน(Interlinked) มาก แสดงว่าชุดข้อมูลนั้นมีความน่าสนใจมากกว่าชุดข้อมูลที่มีการสร้างความสัมพันธ์น้อย ใช้ voiD (Vocabulary of Interlined Dataset) เป็นคำศัพท์ในการอธิบายสารสนเทศ เนื่องจากสามารถจัดเก็บความสัมพันธ์ระหว่างชุดข้อมูลได้ ทำให้สามารถวิเคราะห์และจัดอันดับของชุดข้อมูลจากข้อมูลดังกล่าวได้ ทำการประเมินประสิทธิภาพโดยการเปรียบเทียบผลการจัดลำดับกับวิธีการจัดอันดับข้อมูลวิธีอื่น ผลการวิจัยพบว่า วิธีการวิเคราะห์จำนวนความสัมพันธ์ช่วยให้รายการสารสนเทศที่มีความเกี่ยวข้องสูง ได้แสดงผลในอันดับต้น ๆ

เดลบรูและคณะ (Delbru et al., 2011) ได้นำเสนอวิธีการค้นคืนข้อมูลจากเอกสารแบบกึ่งโครงสร้าง (Semi-Structural) คือ เอกสารอาร์ดีเอฟในโอเพนดาทา การที่กล่าวว่าเป็นกึ่งโครงสร้างเนื่องจากโอเพนดาทามีการกำหนดมาตรฐานสำหรับโครงสร้างเอกสารให้ใช้ภาษาอาร์ดี

เอฟ แต่ไม่มีการกำหนดชุดคำศัพท์และแบบแผนในการอธิบายสารสนเทศ ซึ่งเป็นความท้าทายในการค้นคืนข้อมูลที่มีการเชื่อมโยงกันอยู่ เนื่องจากเอกสารในลิงค์โอเพนเทดามีการใช้คำศัพท์และเค้าร่างเมทาดาทาที่หลากหลาย การค้นคืนข้อมูลที่มีโครงสร้างแตกต่างกันจะต้องทำการปรับชุดเมทาดาทาให้เหมือนกันก่อนจึงจะสามารถค้นคืนได้ งานวิจัยนี้นำเสนอวิธีการกำหนดดัชนีให้กับข้อมูลที่จัดเก็บในโครงสร้างแบบอาร์ดีเอฟ และทำการกำหนดโครงสร้างในแบบของทรี (Tree) เพื่อกำหนดวิธีการตรวจสอบโครงสร้างในการเข้าค้นหาข้อมูลแบบของทรี (Tree Traveling) เพื่อสามารถเข้าถึงข้อมูลได้อย่างรวดเร็วและมีประสิทธิภาพกว่า ผลการวิจัยพบว่า วิธีการกำหนดดัชนีแบบซับซ้อนจะช่วยให้เข้าถึงข้อมูลได้รวดเร็วและได้ข้อมูลที่ตรงกับคำค้นมากกว่าการค้นคืนด้วยวิธีอื่น

เจา (Zhao, 2012) ได้นำเสนอแนวคิดในการแปลงโครงสร้างเมทาดาทาเพื่อแก้ปัญหาข้อมูลที่มีโครงสร้างต่างกันในเครือข่ายข้อมูลโอเพนดาทา โดยมีแนวคิดว่าเอกสารจากโอเพนดาทาเป็นเอกสารที่มีโครงสร้างแบบออนโทโลยีอยู่แล้ว (RDFs) หากเอกสารมีโครงสร้างที่แตกต่างกัน การใช้เทคนิคการแปลงออนโทโลยี (Ontology Mapping) เพื่อรวมคลาส และคุณสมบัติที่มีความคล้ายคลึงกันเข้าด้วยกัน จะสามารถแก้ปัญหาได้ สุดท้ายจะได้ออนโทโลยีของเอกสารที่มีความสมบูรณ์ และสามารถระบุได้ว่าคลาสและคุณสมบัติที่สำคัญของชุดข้อมูลนั้นคือคลาสใดบ้าง และตรวจสอบความผิดปกติของโครงสร้างเมทาดาทาที่ไม่เหมาะสมได้จากการวิเคราะห์ความสัมพันธ์ระหว่างชุดข้อมูล (Interlined) และคุณสมบัติที่ไม่ถูกต้องตามที่ควรจะเป็นเมื่อใช้เก็บข้อมูลจริง

ฮัสซานและคณะ (Hassan et al., 2012) ได้นำเสนอปัญหาเรื่องความถูกต้องของผลการค้นคืนข้อมูลจากลิงค์โอเพนดาทา โดยงานวิจัยนี้ ให้ผู้ใช้เป็นผู้ประเมินความถูกต้องของผลการค้นคืนจากระบบ เมื่อผู้ใช้ประเมินว่ารายการผลการค้นคืนถูกต้องเหมาะสมตรงตามคำถามในการค้นคืน ระบบจะทำการบันทึกและปรับปรุงผลการค้นคืนให้มีความเหมาะสมกับรายการคำค้นมากขึ้นเรื่อย ๆ โดยระบบมีความสามารถในการเรียนรู้ผลที่ผู้ใช้ประเมิน ทำให้การค้นคืนข้อมูลในครั้งถัดไประบบสามารถนำเสนอข้อมูลที่มีความถูกต้องและตรงความต้องการของผู้ใช้มากขึ้น

โนอายและคณะ (Noia et al., 2012) ได้ยกประเด็นปัญหาจำนวนข้อมูลในลิงค์โอเพนดาทาที่มีปริมาณมาก ทำให้ผลการค้นคืนมีรายละเอียดที่เกี่ยวข้องหลายรายการ แต่ไม่สามารถเลือกได้ว่ารายการใดมีความครบถ้วนหรือตรงกับความต้องการ งานวิจัยนี้ใช้วิธีการเปรียบเทียบผลการค้นคืนกับรายการคำค้นด้วยวิธีปริภูมิเวกเตอร์ (Vector Space) โดยใช้วิธีนำคำสำคัญที่ใส่เป็นคำค้นแปลงให้อยู่ในรูปสมการเส้นตรง และนำผลลัพธ์ที่ได้จากสมการไปเปรียบเทียบกับรายการข้อมูลที่มีอยู่ในชุดข้อมูลทีละรายการ หากรายการใดมีผลลัพธ์ใกล้เคียงกับรายการคำค้น แสดงว่ามีความน่าจะเป็นที่รายการนั้นเป็นเรื่องที่เกี่ยวข้องกับรายการคำค้นที่ใส่เข้ามา วิธีการเปรียบเทียบสามารถใช้ลำดับความสำคัญของคำค้นที่ปรากฏในเอกสารมาเป็นส่วนในการวิเคราะห์ความสอดคล้องหรือใกล้เคียงได้ ทำให้วิธีการเปรียบเทียบมีแนวทางที่ยืดหยุ่นและได้ประสิทธิรูปที่ดี

หวังและคณะ (Wang et al., 2013) ได้นำเสนอรูปแบบการวิเคราะห์ข้อมูลภายในลิงค์โอเพนดาทาด้วยวิธีการแปลงชุดเมทาดาทาแบบหลายกรณี เนื่องจากจำนวนข้อมูลในลิงค์โอเพนดาทามีมากขึ้น โครงสร้างข้อมูลที่แตกต่างกันทำให้การทำงานร่วมกันของชุดเมทาดาทาที่แตกต่าง

กันเกิดปัญหาขึ้น งานวิจัยนี้นำเสนอการเปรียบเทียบชุดเมทาดาทาโดยการใช้การแปลงออนโทโลยีหลายรูปแบบ เช่น การวิเคราะห์จากชื่อ การวิเคราะห์จากคุณสมบัติบางประการ การวิเคราะห์จากการปรากฏของคำสำคัญ เป็นต้น ซึ่งผลการเปรียบเทียบของแต่ละกรณีจะถูกนำมาผสานรวมกันจนได้ชุดเมทาดาทาเพียงชุดเดียว ซึ่งผลที่ได้คือชุดเมทาดาทาที่มีความสมบูรณ์ และประเมินความถูกต้องของออนโทโลยีโดยนำไปพัฒนาเป็นระบบค้นคืนและวัดค่าความแม่นยำ ค่าความระลึกและเอฟเมเชอร์ แสดงให้เห็นว่าวิธีการแปลงออนโทโลยีที่พัฒนาขึ้นช่วยให้การค้นคืนสารสนเทศมีความถูกต้องมากกว่าการค้นคืนโดยตรงจากกลุ่มข้อมูลที่ไม่มีการแปลงออนโทโลยี

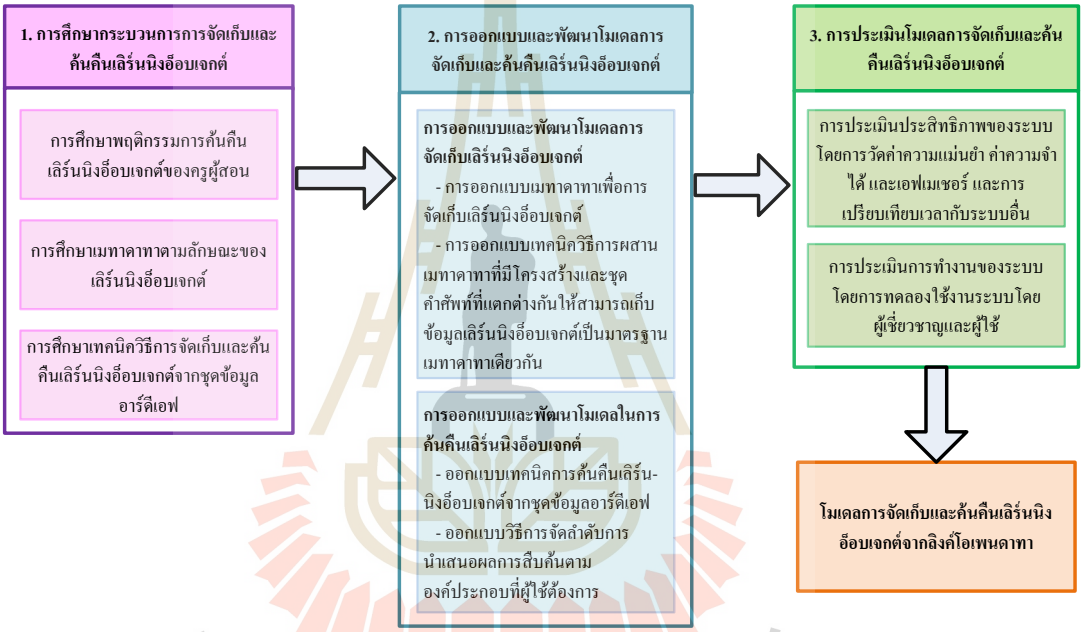
2.5.2 บทสรุปและเปรียบเทียบกระบวนการวิจัย

ตารางที่ 2.7 บทสรุปงานวิจัยที่เกี่ยวข้อง

วิธีการ / ชื่อผู้วิจัย	การเข้าถึงชุดข้อมูล (Dataset Accessing)	เทคนิคการผสานเมทาดาทา (Metadata Mapping)	การจัดลำดับเนื้อหา (Ranking)	รูปแบบคำถาม (Query Processing)
Toupikov, et al., 2009	กำหนดชื่อชุดข้อมูลแบบระบุเจาะจง	-	Page rank	ค้นคืนด้วย SPARQL
Delbru, et al., 2011	กำหนดชื่อชุดข้อมูลแบบระบุเจาะจง	Tree-based Indexing on Entity and Properties	-	ค้นหาแบบ Keyword Search
Zhao, 2012	กำหนดชื่อชุดข้อมูลแบบระบุเจาะจง	Ontology matching on Class	-	ค้นคืนด้วย SPARQL
Hassan, et al., 2012	กำหนดชื่อชุดข้อมูลแบบระบุเจาะจง	User Feedback and Consolidation	-	ค้นคืนแบบ Keyword Search
Noia, et al., 2012	กำหนดชื่อชุดข้อมูลแบบระบุเจาะจง	Vector Space Model on RDF Text-Based	Page Rank	ค้นคืนแบบ Keyword Search
Wang, et al., 2013	กำหนดชื่อชุดข้อมูลแบบระบุเจาะจง	Unified methodology matching	-	ค้นคืนแบบ Keyword Search
งานวิจัยนี้	กำหนดชื่อชุดข้อมูลแบบระบุเจาะจง	Metadata Mapping และ Instance Transformation	Similarity Measurement and Ranking	ค้นคืนแบบ Keyword Search

2.5.3 กรอบแนวคิดในการวิจัย

การวิจัยในครั้งนี้ได้รวบรวมแนวคิด ทฤษฎีในการศึกษาพฤติกรรมการค้นคืนและใช้งานเลิร์นนิงอ็อบเจกต์ของครูผู้สอน เพื่อนำมาออกแบบเมทาดาทาในการจัดเก็บเลิร์นนิงอ็อบเจกต์ รวมถึงเทคนิควิธีแปลงข้อมูลให้อยู่ในมาตรฐานการเชื่อมโยงสู่ลิงค์โอเพนดาทา อีกทั้งการพัฒนาเทคนิคการค้นคืนเลิร์นนิงอ็อบเจกต์ด้วยวิธีการค้นคืนเชิงความหมาย และการจัดลำดับการนำเสนอจากความต้องการของครูผู้สอน เพื่อพัฒนาเป็นโมเดลสำหรับการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากลิงค์โอเพนดาทา โดยกรอบแนวคิดการวิจัยแสดงดังรูปที่ 2.17



รูปที่ 2.17 กรอบแนวคิดการวิจัย

บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อออกแบบ พัฒนาและประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา โดยมุ่งหวังให้ผู้ใช้ซึ่งได้แก่ครูผู้สอนสามารถจัดเก็บเลิร์นนิงอ็อบเจกต์ที่สร้างขึ้น สามารถค้นคืนเลิร์นนิงอ็อบเจกต์ที่เกี่ยวข้องที่นำเข้าจากภายนอกโดยมีโครงสร้างเมทาดาทาที่แตกต่างกันได้โดยง่าย และสามารถจัดลำดับการนำเสนอรายการผลการค้นคืนได้ตามรูปแบบสื่อที่ผู้ใช้ต้องการ ทั้งนี้เพื่อให้ครูผู้สอนสามารถค้นคืนเลิร์นนิงอ็อบเจกต์ในระดับต่าง ๆ ทั้งวัตถุสารสนเทศ วัตถุประยุกต์ และหน่วยบทเรียน โดยสามารถเลือกใช้องค์ประกอบจากสื่อข้อมูลย่อยหรือแอสเซตภายในเลิร์นนิงอ็อบเจกต์ในการผลิตเลิร์นนิงอ็อบเจ็กต์ใหม่ได้ การวิจัยครั้งนี้ใช้รูปแบบของการวิจัยแบบวิจัยและพัฒนา (Research and Development) แบ่งการวิจัยออกเป็น 3 ระยะตามกรอบการวิจัย ได้แก่ (1) การศึกษากระบวนการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ (2) การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา และ (3) การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา โดยแต่ละส่วนมีวิธีดำเนินการวิจัย ดังนี้

3.1 การศึกษากระบวนการการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

การศึกษากระบวนการการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ มีวัตถุประสงค์เพื่อนำผลจากการศึกษามาใช้เป็นข้อมูลเบื้องต้นในการออกแบบเมทาดาทาที่เหมาะสมสำหรับเลิร์นนิงออบเจกต์ และการออกแบบโมเดลในการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ แบ่งการศึกษาออกเป็น 3 ส่วน ได้แก่ (1) การศึกษาลักษณะของเลิร์นนิงอ็อบเจกต์ (2) การศึกษาเทคนิควิธีการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา และ (3) การศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน โดยมีรายละเอียด ดังนี้

3.1.1 การศึกษาลักษณะของเลิร์นนิงอ็อบเจกต์

การศึกษาลักษณะของเลิร์นนิงอ็อบเจกต์มีวัตถุประสงค์เพื่อทราบองค์ประกอบของเลิร์นนิงออบเจกต์ ทั้งในด้านโครงสร้างทางกายภาพและโครงสร้างด้านเนื้อหา ทำการศึกษาโดยการวิเคราะห์แนวคิดทฤษฎีที่เกี่ยวข้องกับการจำแนกประเภทเลิร์นนิงอ็อบเจกต์ และเค้าร่างเมทาดาทาที่ใช้ในการบรรยายเลิร์นนิงอ็อบเจกต์ที่นิยมใช้ในปัจจุบัน โดยนำเสนอไว้ในบทที่ 2 (หัวข้อที่ 2.2 เลิร์นนิงอ็อบเจกต์ และ หัวข้อ 2.3.1 เมทาดาทา หน้า 23-38)

3.1.2 การศึกษาเทคนิควิธีการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา

การศึกษาเทคนิควิธีการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูล โอเพนดาทา มีวัตถุประสงค์เพื่อทราบเทคนิควิธีการในปัจจุบันที่ใช้ในการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา เพื่อนำไปใช้ออกแบบโครงสร้างและองค์ประกอบของโมเดลที่จะพัฒนาขึ้น ทำการศึกษาโดยการวิเคราะห์แนวคิดทฤษฎีที่เกี่ยวข้อง โดยนำเสนอไว้ในบทที่ 2 หัวข้อที่ 2.3.2 คลังคำศัพท์สำหรับการสืบค้นสารสนเทศ (หน้า 40-46) หัวข้อที่ 2.3.4 การผสานเมทาดาทา (หน้า 49-55) และหัวข้อที่ 2.4.1 ภาษาและเครื่องมือสำหรับสืบค้นข้อมูลโอเพนดาทา (หน้า 56-62)

3.1.3 การศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน

การศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอนมีวัตถุประสงค์เพื่อทราบพฤติกรรมการสืบค้นเลิร์นนิงออบเจกต์จากอินเทอร์เน็ต (ข้อมูลที่ใช้ในการสืบค้น การจำแนก และการตัดสินใจเลือกรายการเลิร์นนิงอ็อบเจกต์) ลักษณะของเลิร์นนิงอ็อบเจกต์ที่ต้องการ และการนำเลิร์นนิงอ็อบเจกต์ไปใช้ในการผลิตสื่อการเรียนรู้ ทำการศึกษาโดยการวิจัยเชิงสำรวจ รวบรวมข้อมูลโดยการใช้แบบสอบถาม

3.1.3.1 วิธีดำเนินการวิจัย

การศึกษาพฤติกรรมการค้นคืนและใช้งานเลิร์นนิงอ็อบเจกต์ที่ได้จากอินเทอร์เน็ตบนอินเทอร์เน็ตของครูผู้สอนรายวิชาคณิตศาสตร์หลักสูตรสองภาษา ในโรงเรียนมัธยมศึกษาตอนปลาย ใช้วิธีการวิจัยเชิงสำรวจ เก็บรวบรวมข้อมูลโดยการใช้แบบสอบถาม

3.1.3.2 ประชากรและกลุ่มตัวอย่างในการสำรวจพฤติกรรมการค้นคืน

ประชากรที่ใช้ในการศึกษาครั้งนี้ คือครูผู้สอนรายวิชาคณิตศาสตร์ระดับมัธยมศึกษา ในสังกัดกระทรวงศึกษาธิการ โดยเป็นโรงเรียนที่เข้าโครงการการจัดการเรียนการสอนเป็นภาษาอังกฤษตามที่กระทรวงศึกษาธิการประกาศไว้ ประกอบไปด้วยโรงเรียนสอนหลักสูตรสองภาษาในหลักสูตรอิงลิชโปรแกรม (English Program) และมินิอิงลิชโปรแกรม (Mini English Program) โรงเรียนมาตรฐานสากล (World Class Standard School) ซึ่งมีทั้งโรงเรียนภายใต้การควบคุมดูแลของรัฐ จำนวน 100 แห่ง และโรงเรียนเอกชนจำนวน 50 แห่ง (สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน, 2555) รวมจำนวนโรงเรียนที่มีการเรียนการสอนเป็นภาษาอังกฤษในรายวิชาคณิตศาสตร์ จำนวน 150 โรงเรียน ทั้งนี้จำนวนครูผู้สอนในแต่ละโรงเรียนจะมีจำนวนแตกต่างกันตามจำนวนนักเรียนและจำนวนชั้นเรียนที่เปิดสอน อย่างไรก็ดี ตามนโยบายการจัดหลักสูตรการเรียนสองภาษาที่กำหนดให้แต่ละชั้นเรียนหรือแต่ละห้องเรียนต้องมีครูผู้สอนอย่างน้อย 2 คน (Team Teaching) และอาจเพิ่มจำนวนผู้สอนได้ตามรายวิชาเฉพาะทางอื่น ๆ ซึ่งโรงเรียน

สามารถเปิดสอนได้หนึ่งห้องเรียนต่อระดับชั้นต่อปีการศึกษา (สำนักงานข้าราชการครูและบุคลากรทางการศึกษา, 2551) ดังนั้น คำนวณกลุ่มประชากรจาก 150 โรงเรียน โรงเรียนละ 2 คน จะได้จำนวนประชากรทั้งสิ้น 300 คน โดยผู้วิจัยได้แจกแบบสอบถามเพื่อสำรวจข้อมูลทั้ง 300 คน ซึ่งได้แบบสอบถามกลับคืนทั้งหมด 195 ชุด คิดเป็นร้อยละ 65 ของกลุ่มประชากร ซึ่งเพียงพอต่อความน่าเชื่อถือของข้อมูลจากการประมวลผลแบบสอบถาม ซึ่งปกติจากการคำนวณจำนวนชุดแบบสอบถามที่ส่งกลับคืน ควรมีจำนวนร้อยละ 40 ถึง 50 ของจากกลุ่มประชากรที่มีจำนวนไม่ถึงหนึ่งพันคน (บุญชม ศรีสะอาด, 2543: 38)

3.1.3.4 เครื่องมือในการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน

เครื่องมือที่ใช้ในการรวบรวมข้อมูลพฤติกรรมของครูผู้สอนรายวิชาคณิตศาสตร์ระดับมัธยมศึกษาที่เคยทำการค้นคืนเลิร์นนิงอ็อบเจกต์ เพื่อใช้ประกอบการเรียนการสอนหรือนำมาประกอบการสร้างเลิร์นนิงอ็อบเจกต์ของผู้สอนเอง ได้แก่ แบบสอบถาม (Questionnaire) ที่ผู้วิจัยพัฒนาขึ้นทั้งหมด 4 ตอน ประกอบด้วย

ตอนที่ 1 ข้อมูลพื้นฐานของผู้ตอบแบบสอบถามเพื่อใช้ในการอ้างอิงแหล่งข้อมูลในการวิจัยครั้งนี้ ได้แก่ ตำแหน่งผู้สอน ระดับอาจารย์ ประเภทโรงเรียนที่สังกัด ประสบการณ์การสอนระดับชั้น และหลักสูตรที่รับผิดชอบ เป็นคำถามแบบเลือกตอบ (Check list)

ตอนที่ 2 การเลือกใช้เครื่องมือและแหล่งค้นคืนสารสนเทศบนอินเทอร์เน็ต เพื่อค้นหาข้อมูลเลิร์นนิงอ็อบเจกต์เกี่ยวกับรายวิชาคณิตศาสตร์ เป็นคำถามแบบจัดอันดับ (Ranking Question) และคำถามแบบมาตราส่วนประมาณค่า 5 ระดับ (Rating scale) ตามรูปแบบของ Likert's scale ซึ่งจะมีให้เลือกตอบได้เพียงคำตอบเดียว

ประมาณค่า 5 ระดับ (Rating scale) โดยแต่ละระดับแสดงความความสำคัญที่แตกต่างกันอย่างมีนัยสำคัญ ตามรูปแบบของมาตรวัดไลเคิร์ท Likert's scale ซึ่งผู้ใช้จะต้องทำการเลือกคำตอบได้เพียงคำตอบเดียว โดยมีการกำหนดเกณฑ์การให้คะแนนระดับความสำคัญของปัญหาที่พบในการค้นคืนสารสนเทศบนอินเทอร์เน็ต ดังนี้

ระดับความสำคัญของปัญหาที่พบในการค้นคืน	ระดับคะแนน
เป็นปัญหามากที่สุด	5
เป็นปัญหามาก	4
เป็นปัญหาปานกลาง	3
เป็นปัญหาน้อย	2
เป็นปัญหาน้อยที่สุด	1

ตอนที่ 3 การเลือกใช้ข้อมูลจากแหล่งสารสนเทศบนอินเทอร์เน็ต เพื่อนำมาใช้ในการสร้างสื่อการเรียนรู้ การนำไปประกอบการเรียนการสอน เป็นคำถามแบบมาตราส่วนประมาณค่า 5 ระดับ (Rating scale) โดยแต่ละระดับแสดงความความสำคัญที่แตกต่างกันอย่างมีนัยสำคัญ ตามรูปแบบของมาตรวัดไลเคิร์ท Likert's scale ซึ่งผู้ใช้จะต้องทำการเลือกคำตอบได้เพียงคำตอบเดียวที่ตนเองเห็นว่าตรงกับความต้องการหรือรู้สึกมากที่สุด

คำถามแบบมาตราส่วนประมาณค่า 5 ระดับ (Rating scale) ตามรูปแบบของ Likert's scale มีการกำหนดเกณฑ์การให้คะแนนระดับการให้ความสำคัญกับเขตข้อมูลย่อยที่ใช้อธิบายเลิร์นนิงอ็อบเจกต์ ดังนี้

ระดับการให้ความสำคัญในการพิจารณาเลือกใช้ข้อมูล	ระดับคะแนน
ให้ความสำคัญมากที่สุด	5
ให้ความสำคัญมาก	4
ให้ความสำคัญปานกลาง	3
ให้ความสำคัญน้อย	2
ให้ความสำคัญน้อยที่สุด	1

วิธีการแปลผลแบบสอบถามส่วนนี้ได้ใช้ค่าเฉลี่ย ตามเกณฑ์คะแนน ของแบบสอบถามตอนที่ 2 และ ตอนที่ 3 ดังนี้

(คะแนนสูงสุด – คะแนนต่ำสุด)/จำนวนชั้น

ตอนที่ 4 การจัดเก็บเลิร์นนิงอ็อบเจกต์ในสภาพปัจจุบัน และความต้องการระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ แบบจัดอันดับ (Raking Question)

3.1.3.5 การวัดประสิทธิภาพของเครื่องมือที่ใช้การสำรวจพฤติกรรมการค้นคืน

การหาประสิทธิภาพของแบบสอบถาม ผู้วิจัยได้ทดสอบคุณภาพของเครื่องมือ 2 ส่วน คือการทดสอบความตรง (Validity) และความเชื่อถือได้ (Reliability) ของแบบสอบถามก่อนนำไปเก็บข้อมูลจริง ดังนี้ คือ

3.1.3.5.1 การตรวจสอบความตรง (Validity) โดยการนำแบบสอบถามให้ผู้เชี่ยวชาญ พิจารณาในด้านความเที่ยงตรง ความเหมาะสมของเนื้อหา ความเหมาะสมในการใช้ภาษา และความชัดเจนของคำถามในแบบสอบถาม เพื่อให้แบบสอบถามสามารถสื่อความหมายได้ตรงตามความประสงค์ของผู้วิจัย และผู้ตอบแบบสอบถาม ซึ่งจะให้ผู้ทรงคุณวุฒิที่มีความรู้ความ

เชี่ยวชาญกับสิ่งที่จะวัดเป็นผู้ทำการตรวจสอบ สำหรับงานวิจัยโดยทั่วไปจะมีจำนวนผู้ทรงคุณวุฒิร่วมตรวจสอบคุณภาพของเครื่องมือวัด 3 คน ประกอบด้วย

- ผู้ช่วยศาสตราจารย์ ดร.กรองทิพย์ นาควิเชตร รองอธิการบดีฝ่ายวิชาการมหาวิทยาลัยวงษ์ชวลิตกุล (ผู้เชี่ยวชาญด้านการวิจัย และสถิติ)
- อาจารย์ ดร.บุษยา วงษ์ชวลิตกุล มหาวิทยาลัยวงษ์ชวลิตกุล (ผู้เชี่ยวชาญด้านการวิจัย)
- ดร.บัณฑิต วรรณประพันธ์ นักวิชาการสาธารณสุขชำนาญการพิเศษ สำนักงานป้องกันและควบคุมโรค ที่ 9 นครราชสีมา (ผู้เชี่ยวชาญด้านการวิจัย และสถิติ)

โดยการให้คะแนนรายข้อของผู้เชี่ยวชาญจะมี 3 ค่า คือ

ค่า +1 คือ ผู้ตรวจสอบแน่ใจว่าข้อคำถามนั้นสามารถใช้วัดค่าตัวแปรที่จะศึกษาได้

ค่า 0 คือ ผู้ตรวจสอบไม่แน่ใจว่าข้อคำถามนั้นสามารถใช้วัดค่าตัวแปรที่ศึกษาได้หรือไม่

ค่า -1 คือ ผู้ตรวจสอบแน่ใจว่าข้อคำถามนั้นไม่สามารถใช้วัดค่าตัวแปรที่จะศึกษาได้

การคำนวณหาค่าดัชนีความสอดคล้อง (Index of consistency: IOC) โดยเลือกเฉพาะข้อคำถามที่มีค่าดัชนีความสอดคล้องตั้งแต่ 0.5 เป็นต้นไป เพื่อให้ได้ข้อคำถามที่มีคุณภาพ แล้วจึงนำไปปรับปรุงแก้ไข โดยการทดสอบหาความเที่ยงตรงของแบบสอบถามปัจจัยการเลือกใช้สารสนเทศบนอินเทอร์เน็ตใช้สูตรดัชนีค่าความสอดคล้อง IOC (กรมวิชาการ กระทรวงศึกษาธิการ, 2545: 65)

$$IOC = \frac{\Sigma R}{N} \qquad (3\text{-}1)$$

เมื่อ

IOC แทน ดัชนีความสอดคล้องระหว่างจุดประสงค์กับเนื้อหาหรือระหว่างข้อสอบกับจุดประสงค์

ΣR แทน ผลรวมระหว่างคะแนนความคิดเห็นของผู้เชี่ยวชาญทั้งหมด

N แทน จำนวนผู้เชี่ยวชาญ

3.1.3.5.2 การทดสอบความเชื่อถือได้ (Reliability) ของแบบสอบถาม โดยการนำแบบสอบถามที่ได้รับการตรวจสอบและปรับปรุงแล้วไปทดลองใช้ (Try out) กับครูคณิตศาสตร์ที่ทำการสืบค้นเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตที่ไม่ใช่กลุ่มตัวอย่าง จำนวน 30 คน และนำผลที่ได้มาคำนวณหาค่าความเชื่อมั่นด้วยวิธีการหาค่าสัมประสิทธิ์อัลฟ่า (Alpha Coefficient) ของครอนบาค (Cronbach, 1970) มีสูตรดังนี้

$$\alpha = \frac{K}{K-1}\left\{1 - \frac{\sum s_i^2}{s_t^2}\right\} \qquad (3\text{-}2)$$

โดยที่ α = ค่าสัมประสิทธิ์ความเชื่อมั่น

K = จำนวนข้อคำถามของแบบสอบถาม

$\sum s_i^2$ = ผลรวมของความแปรปรวนของคะแนนแต่ละข้อ

S_t^2 = ความแปรปรวนของคะแนนรวม

ผลการทดสอบความเชื่อถือได้ของแบบสอบถามการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน เท่ากับ 0.834 ค่าสัมประสิทธิ์อัลฟาที่ได้มีค่าเกิน 0.70 ซึ่งถือว่ายอมรับได้จึงดำเนินการเก็บข้อมูลต่อไป

3.1.3.6 การวิเคราะห์และแปลผลข้อมูลการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์

การวิเคราะห์ข้อมูลจากแบบสอบถามการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอนรายวิชาคณิตศาสตร์ ประกอบด้วยข้อมูล 4 ตอน ข้อมูลที่รวบรวมได้จะนำไปวิเคราะห์โดยใช้โปรแกรมสำเร็จรูปเอสพีเอสเอส (SPSS) และนำเสนอในรูปตารางประกอบการแปลความหมายเชิงบรรยาย เพื่ออธิบายข้อมูลในด้านต่าง ๆ ดังนี้

ตอนที่ 1 ข้อมูลพื้นฐานของผู้ตอบแบบสอบถามเพื่อใช้ในการอ้างอิงแหล่งข้อมูลในการวิจัยครั้งนี้ ได้แก่ ตำแหน่งผู้สอน ระดับอาจารย์ ประเภทโรงเรียนที่สังกัด ประสบการณ์การสอน ระดับชั้นและหลักสูตรที่รับผิดชอบ วิเคราะห์ข้อมูลโดยการแจกแจงค่าความถี่ (Frequencies) การหาค่าสัดส่วนหรือร้อยละ (Percentage)

ตอนที่ 2 การเลือกใช้เครื่องมือและแหล่งค้นคืนสารสนเทศบนอินเทอร์เน็ต เพื่อค้นหาข้อมูลเลิร์นนิงอ็อบเจกต์เกี่ยวกับรายวิชาคณิตศาสตร์ วิเคราะห์ข้อมูลโดยการแจกแจงค่าความถี่ (Frequencies) การหาค่าสัดส่วนหรือร้อยละ (Percentage) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: S.D.)

ตอนที่ 3 การเลือกใช้ข้อมูลจากแหล่งสารสนเทศบนอินเทอร์เน็ต เพื่อนำมาใช้ในการสร้างสื่อการเรียนรู้ การนำไปประกอบการเรียนการสอน เป็นคำถามแบบมาตราส่วนประมาณค่า 5 ระดับ (Rating scale) ตามรูปแบบของ Likert's scale ซึ่งจะมีให้เลือกตอบได้เพียงคำตอบเดียว วิเคราะห์ วิเคราะห์ข้อมูลโดยการแจกแจงค่าความถี่ (Frequencies) การหาค่าสัดส่วนหรือร้อยละ (Percentage) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: S.D.)

ตอนที่ 4 การจัดเก็บเลิร์นนิงอ็อบเจกต์ในสภาพปัจจุบัน และความต้องการระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ แบบจัดอันดับ (Raking Question) วิเคราะห์ข้อมูลโดยการแจกแจงค่าความถี่ (Frequencies) การหาค่าสัดส่วนหรือร้อยละ (Percentage) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: S.D.)

การวิเคราะห์ข้อมูลส่วนคำถามปลายเปิดเกี่ยวกับความคิดเห็นและข้อเสนอแนะใช้การวิเคราะห์เนื้อหาโดยวิธีสรุปประเด็นเนื้อหา และจัดแยกเป็นหมวดหมู่

3.2 การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา

การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ สามารถแบ่งกระบวนการทำงานของโมเดลออกเป็น 2 ส่วนได้แก่ 3.2.1 การออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ มีจุดประสงค์เพื่อออกแบบชุดเมทาดาทาที่จะนำมาจัดเก็บข้อมูลของเลิร์นนิงอ็อบเจกต์ให้ตรงกับความต้องการของครูผู้สอน และ 3.2.2 การออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ มีวัตถุประสงค์เพื่อออกแบบเทคนิควิธีการค้นคืนข้อมูลเมทาดาทาและไฟล์เลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา พร้อมทั้งการจัดลำดับการนำเสนอผลการสืบค้นตามที่ผู้ใช้ต้องการ โดยจะอธิบายในแต่ละส่วนตามลำดับ ดังนี้

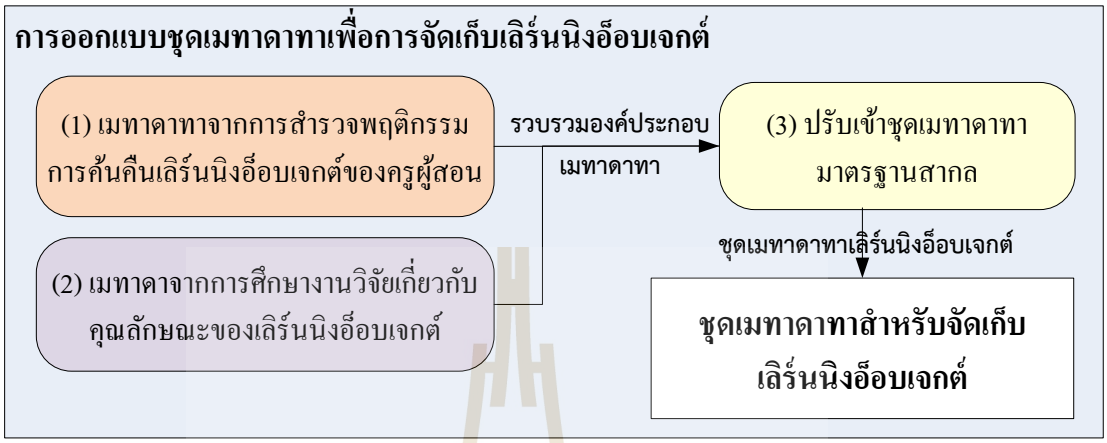
3.2.1 การออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์

ในกระบวนการจัดเก็บเลิร์นนิงอ็อบเจกต์จะมีกระบวนการดังต่อไปนี้

3.2.1.1 การออกแบบชุดเมทาดาทาสำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์

ชุดเมทาดาทาที่ใช้อธิบายเลิร์นนิงอ็อบเจกต์ในระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ในงานวิจัยนี้ ออกแบบและพัฒนาขึ้นจากข้อมูลเมทาดาทา 2 ส่วน ได้แก่ (1) เมทาดาทาที่ได้จากการศึกษาพฤติกรรมค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน และ (2) เมทาดาทาที่ศึกษาจากงานวิจัยและข้อกำหนดตามมาตรฐานโครงสร้างด้านกายภาพและเนื้อหาของเลิร์นนิงอ็อบเจกต์ที่ควรมี เพื่อใช้ในการแลกเปลี่ยนและใช้ข้อมูลของส่วนประกอบข้อมูลย่อยภายในเลิร์นนิงอ็อบเจกต์แต่ละรายการเลิร์นนิงอ็อบเจกต์ซ้ำได้ โดยรวบรวมเมทาดาที่ได้ทั้งสองขั้นตอนมาปรับเข้ากับเมทาดาทา

มาตรฐานสากลในขั้นตอนที่ (3) เพื่อกำหนดเป็นชุดเมทาดาทาสำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์ในโมเดลนี้ โดยแสดงดังรูปที่ 3.1



รูปที่ 3.1 กรอบแนวคิดการออกแบบชุดเมทาดาทาสำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์

(1) เมทาดาทาที่ได้จากการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน จากการศึกษาพฤติกรรมการเลือกใช้เมทาดาทาในการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน พบเมทาดาทาที่ผู้ใช้เลือกใช้ในการพิจารณารายละเอียดของเลิร์นนิงอ็อบเจกต์ในระดับ “มาก” ขึ้นไปรวมจำนวน 10 องค์ประกอบ ได้แก่ ชื่อเรื่อง (Title) คำสำคัญ (Keyword) ประเภทของไฟล์ (Format) รายละเอียด (Description) บทคัดย่อ (Abstract) หัวเรื่อง (Subject) ภาษา (Language) แหล่งที่มาของข้อมูล (Source) ระดับความยากง่ายของเนื้อหา (Difficulty) ผู้รับผิดชอบเนื้อหา (Publisher) ซึ่งเมทาดาทาข้างต้นจะนำไปประกอบการกำหนดชุดเมทาดาทาสำหรับอธิบายเลิร์นนิงอ็อบเจกต์ในกระบวนการถัดไป

(2) เมทาดาทาที่ศึกษาจากโครงสร้างด้านกายภาพและเนื้อหาของเลิร์นนิงอ็อบเจกต์ โดยใช้ตามแนวคิดของมาลิสัน (Mallison, 2001) ที่จำแนกองค์ประกอบออกเป็น 3 ระดับ ได้แก่ วัตถุสารสนเทศ (Information Object) วัตถุประยุกต์ (Application Object) และหน่วยบทเรียน (Lesson or Aggregates Assemblies) โดยเลิร์นนิงอ็อบเจกต์แต่ละรายการจะมีสื่อข้อมูลย่อย (Media Data Element) บ้างจะถูกเรียกว่าข้อมูลดิบหรือแอสเซต (Asset) จะถูกนำไปจัดเรียง รวบรวม และสร้างความสัมพันธ์ระหว่างข้อมูลดิบแต่ละชิ้น ให้มีขอบเขตเนื้อหาและวัตถุประสงค์ที่เหมาะสมในการใช้งานในแต่ละระดับ และจะถูกเรียกชื่อแตกต่างออกไปตามขนาดและขอบเขตของเนื้อหาและวัตถุประสงค์ในการเรียนรู้ ซึ่งจากแนวคิดข้างต้นจึงได้ศึกษาคุณลักษณะของเลิร์นนิงอ็อบเจกต์ตามแนวคิดของลอม (LOM) ได้ดังนี้

(2.1) เลิร์นนิงอ็อบเจกต์ต้องมีวัตถุประสงค์ชัดเจน (Objective) เพื่อกำหนดขอบเขตและการนำเสนอของเลิร์นนิงอ็อบเจกต์แต่ละรายการ ผู้สอนนิยมใช้วัตถุประสงค์เป็นกรอบของการสร้างหรือเลือกใช้เลิร์นนิงอ็อบเจกต์ให้ตรงกับเนื้อหาสาระรายวิชาและวัตถุประสงค์แต่ละข้อของบทเรียน ดังนั้น ในเลิร์นนิงอ็อบเจกต์แต่ละรายการสามารถมีวัตถุประสงค์ได้หลายข้อ

(2.2) เลิร์นนิงอ็อบเจกต์ก่อนหน้า (Pre-Requisites) เป็นข้อกำหนดของรายวิชาหรือเนื้อหาในบทเรียนต่าง ๆ ที่บางครั้งจำเป็นจะต้องมีการเรียนหรือศึกษาในเนื้อหาบางเรื่องมาก่อน หรือมีความรู้พื้นฐานในบางรายวิชามาก่อน เป็นต้น เมื่อนำมากำหนดลักษณะการจัดเก็บเลิร์นนิงอ็อบเจกต์จึงต้องนำข้อกำหนดนี้เป็นส่วนประกอบหนึ่งในการกำหนดความสัมพันธ์ของเนื้อหาและรายวิชาแต่ละส่วนตามความเหมาะสมและตามโครงสร้างหลักสูตรที่ปรับเปลี่ยนได้ตลอดเวลา เมทาดาทาที่ใช้เพื่อกำหนดนอกเหนือจากการกำหนดตามเนื้อหาและวัตถุประสงค์ของบทเรียนแล้ว สามารถกำหนดได้จากเมทาดาทาระดับการศึกษา (Educational Level) เพื่อให้ผู้ใช้ทราบว่าเนื้อหานี้สำหรับกลุ่มผู้ใช้ระดับใด หรือต้องผ่านการศึกษาระดับใดมาก่อนที่จะใช้เลิร์นนิงอ็อบเจกต์รายการนี้

(2.3) ลักษณะและระดับการปฏิสัมพันธ์ (Interactive Type and Level) แสดงถึงลักษณะที่ผู้ใช้สามารถมีปฏิสัมพันธ์กับเลิร์นนิงอ็อบเจกต์แต่ละรายการ ในกระบวนการเรียนรู้จากเลิร์นนิงอ็อบเจกต์จะมีองค์ประกอบในการมีปฏิสัมพันธ์กับผู้ใช้ในหลาย ๆ ส่วนตั้งแต่การดู ฟัง หรืออ่านตาม ไปถึงการตอบคำถาม การแสดงวิธีทำ หรือการประเมินผลจากการทดสอบ เป็นต้น ดังนั้นระดับการปฏิสัมพันธ์จะสามารถบ่งบอกถึงระดับความซับซ้อนของเลิร์นนิงอ็อบเจกต์และลักษณะการนำไปใช้ในกลุ่มผู้ใช้แต่ละกลุ่มได้

(2.4) ประเภทของเลิร์นนิงอ็อบเจกต์ (Learning Object Type) ในกรณีนี้หมายถึงการกำหนดลักษณะประเภทเนื้อหาตามข้อกำหนดของไอเอ็มเอส ไอทริเปิลอีลอม (IMS-IEEE LOM) โดยกำหนดประเภทของเลิร์นนิงอ็อบเจกต์ตามลักษณะของเนื้อหา ได้แก่

- แบบฝึกหัด
- แบบจำลองเหตุการณ์
- แบบทดสอบ
- การทดลอง
- โจทย์ปัญหา
- บทบรรยายเนื้อหา
- แบบสอบถาม

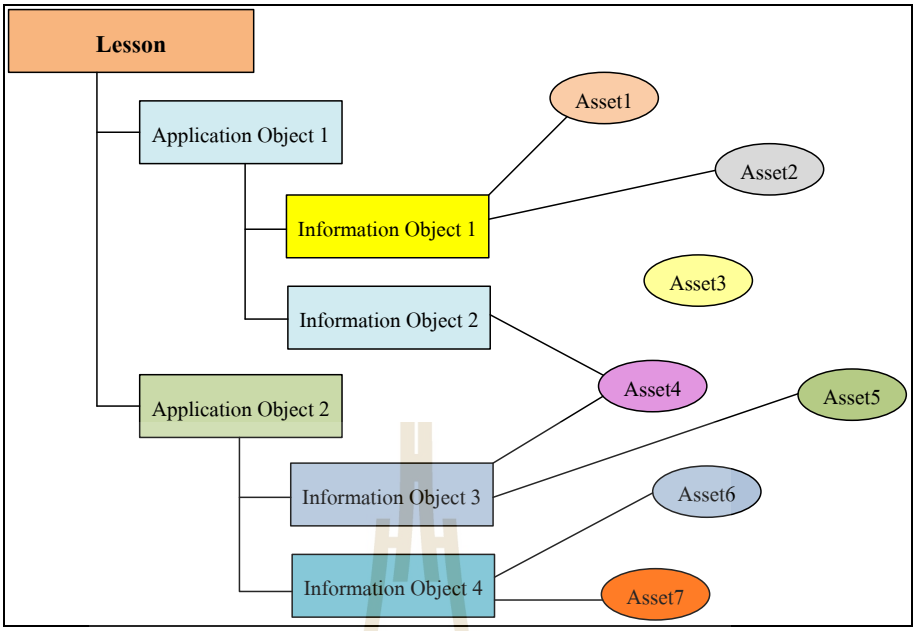
- แผนภาพ รูปภาพ กราฟ ตาราง
- สไลด์นำเสนอ
- แบบการเรียนรู้ด้วยตนเอง

โดยประเภทเนื้อหาของเลิร์นนิงอ็อบเจกต์ข้างต้น จะนำไปใช้ในการออกแบบโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ เพื่อให้ผู้ใช้เลือกประเภทของเนื้อหาได้หลากหลายและตรงกับประเภทของเลิร์นนิงอ็อบเจกต์แต่ละรายการมากที่สุด อีกทั้งช่วยในการจำแนกลักษณะเนื้อหาเพื่อนำไปใช้ในการจัดกลุ่มการนำเสนอผลการค้นคืนให้มีความสอดคล้องกับความต้องการของผู้ใช้

(2.5) ระดับความยากง่ายและระยะเวลาในการเรียน (Difficulty and learning Time) เลิร์นนิงอ็อบเจกต์แต่ละรายการควรมีการกำหนดระดับความยากของเนื้อหา ร่วมกับระยะเวลาในการเรียน เพื่อให้ผู้สอนสามารถเตรียมกระบวนการเรียนการสอนให้สอดคล้องกับระดับเนื้อหา และวิเคราะห์วิธีการนำเสนอให้ผู้เรียนเข้าใจในเนื้อหาโดยง่าย อีกทั้งช่วยในการกำหนดช่วงเวลาเรียนได้ตามระยะเวลาการศึกษาร่วมกับเนื้อหาส่วนอื่นได้อย่างเหมาะสม

ทั้งนี้การกำหนดระดับความยากและระยะเวลาในการเรียน ในระบบต้นแบบจะอนุญาตให้ผู้สอนเป็นผู้กำหนดเองเพื่อให้ระดับเนื้อหาสอดคล้องกับระดับผู้เรียนและลักษณะการสอนมากที่สุด รวมถึงระยะเวลาในการเรียนให้อ้างอิงตามขอบเขตวัตถุประสงค์และระยะเวลาเรียนตามหลักสูตรแกนกลางของกระทรวงศึกษาธิการเป็นหลัก

(2.6) การแลกเปลี่ยนระหว่างเลิร์นนิงอ็อบเจกต์ (Sharable Content Object) ตามแนวคิดการสร้างเลิร์นนิงอ็อบเจกต์จากข้อมูลดิบ (Asset) หลาย ๆ ชิ้น หรือนำเลิร์นนิงอ็อบเจกต์มาแยกส่วนเพื่อนำเนื้อหาบางส่วนหรือเลิร์นนิงอ็อบเจกต์ย่อยบางส่วนมาใช้งาน จึงต้องมีการกำหนดการเชื่อมโยงหรืออธิบายความสัมพันธ์ระหว่างเลิร์นนิงอ็อบเจกต์ เพื่ออธิบายถึงองค์ประกอบภายในเลิร์นนิงอ็อบเจกต์แต่ละรายการว่ามีองค์ประกอบอะไรรวมกันอยู่บ้าง สามารถแสดงได้ดังรูปที่ 3.2



รูปที่ 3.2 โครงสร้างเลิร์นนิงอ็อบเจกต์ที่มีการใช้ข้อมูลย่อยร่วมกัน

จากรูปที่ 3.2 แสดงโครงสร้างบทเรียนหนึ่งหนึ่งที่ประกอบขึ้นจากวัตถุประยุกต์ (Application Object) หลายรายการ โดยภายในวัตถุประยุกต์แต่ละรายการจะมีวัตถุสารสนเทศประกอบอยู่หลายหัวข้อ (Information Object) ตั้งแต่หนึ่งรายการขึ้นไป ซึ่งวัตถุสารสนเทศบางรายการสามารถแยกรายละเอียดเป็นข้อมูลย่อย (Asset) แต่ละชิ้นได้ เช่น ข้อความเนื้อหา ไฟล์เสียงดนตรี ภาพเคลื่อนไหว หรือรูปภาพนิ่ง เป็นต้น แต่บางวัตถุสารสนเทศมีรายละเอียดข้อมูลย่อยที่ถูกบีบอัดรวมเข้าเป็นเนื้อเดียวกันและใช้งานพร้อมกัน จึงไม่สามารถแยกชิ้นข้อมูลย่อยออกจากวัตถุสารสนเทศรายการหลักได้ เช่น ไฟล์วีดีโอ ไฟล์เกม สื่อคอมพิวเตอร์ช่วยสอนสำเร็จรูป เป็นต้น จากการศึกษาถึงข้อมูลคุณสมบัติ ของเลิร์นนิงอ็อบเจกต์จากงานวิจัยข้างต้น จะได้องค์ประกอบเมทาดาทาเพิ่มเติมในชุดโครงสร้างอีก 7 องค์ประกอบ ได้แก่ ระดับการศึกษาหรือกลุ่มเป้าหมาย (EducationLevel/Audience) รูปแบบการปฏิสัมพันธ์ (Interactive Type) ระดับการปฏิสัมพันธ์ (Interactive Level) ประเภทเลิร์นนิงอ็อบเจกต์ (Learning object Type) ระยะเวลาการเรียน (Learning Time) มีส่วนประกอบ (hasPart) เป็นส่วนประกอบของ (isPartOf) เมื่อรวบรวมเมทาดาทาจากทั้ง 2 ส่วนจะได้เป็นชุดเมทาดาทาที่ใช้อธิบายเลิร์นนิงอ็อบเจกต์ในงานวิจัยนี้ทั้งสิ้น 17 องค์ประกอบ

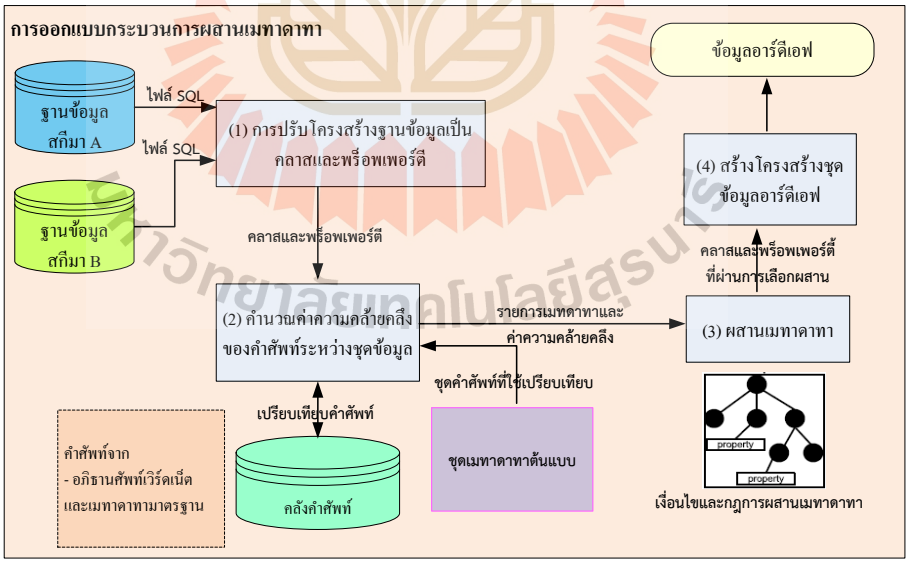
(3) การปรับเข้าชุดเมทาดาทามาตรฐานสากล

การออกแบบชุดเมทาดาทาเพื่อจัดเก็บเลิร์นนิงอ็อบเจกต์ จากผลการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ การศึกษาลักษณะโครงสร้างเนื้อหาและโครงสร้างทางกายภาพ

และการศึกษาเมทาดาทาที่ใช้ในการบรรยายเลิร์นนิงอ็อบเจกต์ที่มีอยู่ในปัจจุบัน ได้แก่ มาตรฐาน Dublin Core มาตรฐาน และ IEEE LOM ข้อมูลทั้ง 3 ส่วนมาวิเคราะห์และกำหนดชุดขององค์ประกอบเมทาดาทาที่สามารถบรรยายลักษณะทางบรรณานุกรมเพื่อการสืบค้นเลิร์นนิงอ็อบเจกต์ (เช่น ชื่อเรื่อง ชื่อผู้แต่ง ปีที่ผลิต หัวเรื่อง ระดับการศึกษา) สามารถบรรยายโครงสร้างทางกายภาพ (เช่น วัตถุสารสนเทศ วัตถุประยุกต์ หน่วยบทเรียน สื่อข้อมูลย่อย) และสามารถบรรยายเนื้อหาของเลิร์นนิงอ็อบเจกต์ (เช่น บทเรียน ข้อสอบ กิจกรรม) เพื่อสนับสนุนการเชื่อมโยงข้อมูลย่อยภายในเลิร์นนิงอ็อบเจกต์ และกำหนดให้เป็นเมทาดาทาสำหรับอธิบายเลิร์นนิงอ็อบเจกต์ในงานวิจัยครั้งนี้ โดยรายละเอียดเมทาดาทาจะนำเสนอในบทที่ 4 ส่วนสรุปและอภิปรายผล

3.2.1.2 การออกแบบกระบวนการผสานเมทาดาทา

การออกแบบกระบวนการผสานเมทาดาทา จะใช้แนวคิดทฤษฎีในการแปลงตารางข้อมูลให้อยู่ในรูปแบบของคลาสและพร็อพเพอร์ตี เพื่อนำมาเปรียบเทียบความคล้ายคลึงของชุดคำศัพท์ระหว่างชุดเมทาดาทามาตรฐานที่กำหนดไว้กับชุดเมทาดาทามาตรฐานอื่น ๆ จากนั้นจึงจะนำค่าความคล้ายคลึงไปทำการวิเคราะห์ในการกำหนดวิธีการผสานเมทาดาทาให้เป็นมาตรฐานเมทาดาทารูปแบบเดียวกัน ก่อนจะกำหนดความสัมพันธ์ให้กับคลาสและพร็อพเพอร์ตีให้เป็นโครงสร้างอาร์ดีเอฟแล้วจึงจัดเก็บในคลังข้อมูล ดังรูปที่ 3.3



รูปที่ 3.3 กรอบแนวคิดการผสานเมทาดาทาและแปลงโครงสร้างเป็นอาร์ดีเอฟ

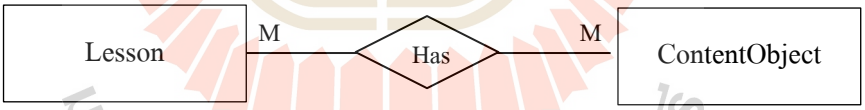
กระบวนการผสานเมทาดาทาในงานวิจัยนี้ออกแบบกระบวนการแยกเป็น 4 ขั้นตอน ประกอบด้วย (1) การปรับโครงสร้างฐานข้อมูลเป็นคลาสและพร็อพเพอร์ตี (2) การคำนวณ

ค่าความคล้ายคลึงของคำศัพท์ระหว่างชุดข้อมูล (3) การผสานเมทาดาทา และ (4) การสร้างชุดข้อมูลอาร์ดีเอฟ และส่งไปจัดเก็บในคลังข้อมูล ทั้งนี้การแปลงและผสานเมทาดาทามีจุดประสงค์เพื่อรองรับการนำเข้าข้อมูลจากหลายแหล่งที่มีมาตรฐานเมทาดาทาที่แตกต่างกันให้เก็บข้อมูลร่วมกันได้ โดยมีรายละเอียดดังนี้

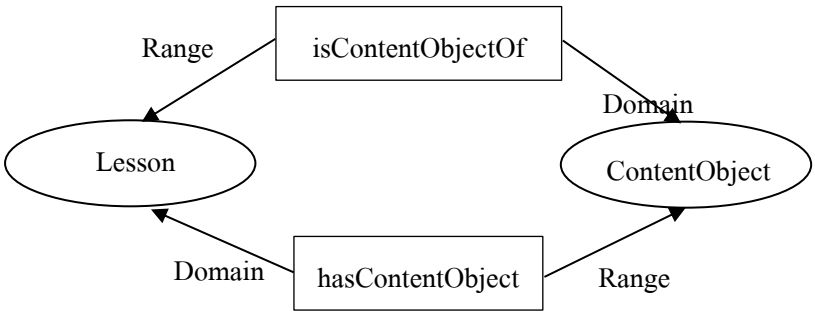
(1) การปรับโครงสร้างฐานข้อมูลเป็นคลาสและพร็อพเพอร์ตี

ในขั้นตอนแรกข้อมูลที่นำเข้าสู่กระบวนการผสานเมทาดาทาจะถูกปรับเปลี่ยนโครงสร้างจากข้อมูลที่เป็นความสัมพันธ์แบบตารางข้อมูลเชิงสัมพันธ์ให้อยู่ในรูปแบบของคลาส (Class) และคุณสมบัติของคลาส (Data Properties)โดยการดึงชื่อตารางแต่ละตารางพร้อมทั้งแอตทริบิวต์ภายในตารางออกมา และทำการปรับให้เป็นคลาสและคุณสมบัติของคลาส และในส่วนความสัมพันธ์ระหว่างตารางก็จะมีการเปลี่ยนรูปแบบให้เป็นความสัมพันธ์ระหว่างคลาส (Object Properties) โดยสามารถจำแนกรูปแบบได้ดังต่อไปนี้

(1.1) กรณีความสัมพันธ์แบบหลายต่อหลาย (Many to Many: M:M) (Cullot, Ghawi and Yetongnon, 2007) ดังรูปที่ 3.4 เมื่อเกิดความสัมพันธ์ดังกล่าว จะเกิดอ็อบเจกต์พร็อพเพอร์ตีระหว่างสองตารางโดยตารางทั้งสองจะถูกสร้างเป็นคลาส และคุณสมบัติระหว่างคลาสจะเป็นส่วนกลับความสัมพันธ์ (inversProperty) ดังคุณสมบัติ hasContentObject และ isContentObjectof ดังรูปที่ 3.5

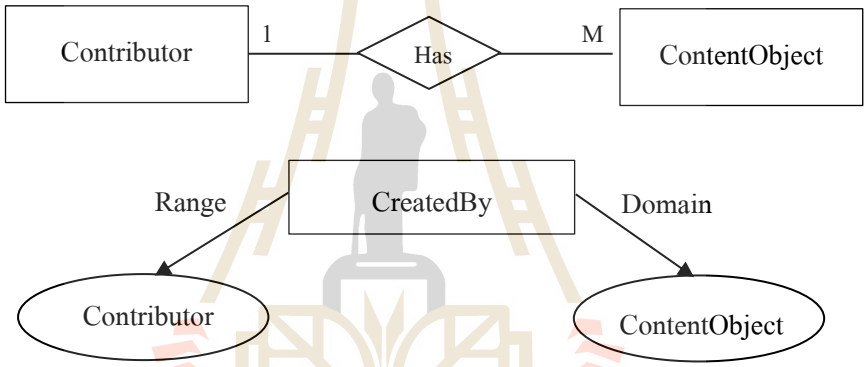


รูปที่ 3.4 ความสัมพันธ์ระหว่างบทเรียนและเลิร์นนิงอ็อบเจกต์

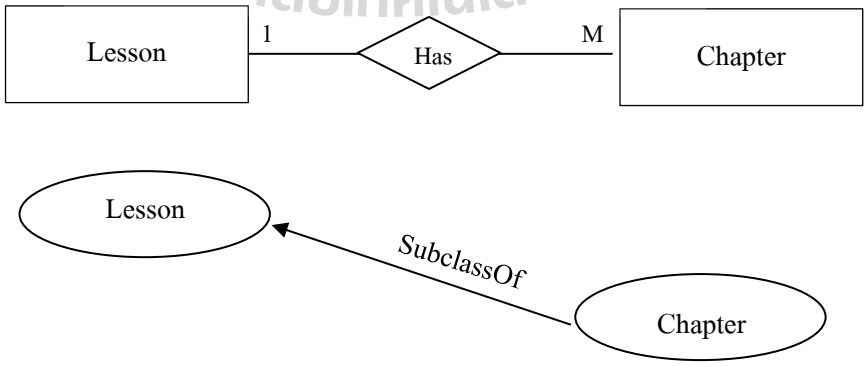


รูปที่ 3.5 ความสัมพันธ์ระหว่างคลาสบทเรียนและคลาสเลิร์นนิงอ็อบเจกต์

(1.2) กรณีความสัมพันธ์แบบหนึ่งต่อหลาย (One to Many: 1:M) ความสัมพันธ์ของตารางในรูปแบบนี้สามารถแปลงได้ 2 ลักษณะ คือ 1) กรณีที่ตารางทั้งสองไม่ใช่ข้อมูลที่เป็นซับเซตกันหรือเป็นข้อมูลคนละประเภท จะทำการสร้างอ็อบเจกต์พร็อพเพอร์ตีระหว่างคลาสทั้งสอง โดยกำหนดให้เรนจ์ (Range) ของอ็อบเจกต์พร็อพเพอร์ตีอยู่ทางคลาสที่มาจากตารางที่มีความสัมพันธ์แบบหนึ่ง ส่วนโดเมน (Domain) จะอยู่ทางคลาสที่มาจากตารางที่มีความสัมพันธ์แบบหลาย ดังรูปที่ 3.6 และ 2) กรณีที่ตารางทั้งสองเป็นข้อมูลที่เป็นซับเซตกัน หรือมีข้อมูลในกลุ่มเดียวกัน จะทำการกำหนดความเป็นคลาสหลักและคลาสย่อย (Class and SubClass) ระหว่างตารางทั้งสองเมื่อเปลี่ยนตารางเป็นคลาส คลาสที่มีความสัมพันธ์แบบหลาย (Many) จะเป็นคลาสย่อย (SubClass) ของคลาสที่มีความสันพันธ์แบบหนึ่ง (One) ดังรูปที่ 3.7



รูปที่ 3.6 การแปลงความสัมพันธ์ของตารางความสัมพันธ์แบบ 1 tom Many แบบที่ 1



รูปที่ 3.7 การแปลงความสัมพันธ์ของตารางความสัมพันธ์แบบ 1 to Many แบบที่ 2

มหาวิทยาลัยเทคโนโลยีสุรนารี

จากข้อกำหนดการแปลงความสัมพันธ์ระหว่างตารางเพื่อแปลงเป็นความสัมพันธ์ระหว่างคลาสและคุณสมบัติของคลาส (Data Properties) และความสัมพันธ์ระหว่างคลาสต่าง ๆ (Object Properties) ข้างต้น ก็จะได้องค์ประกอบเพื่อนำไปสร้างเป็นอาร์ดีเอฟต่อไป

(2) การคำนวณค่าความคล้ายคลึงของคำศัพท์ระหว่างชุดข้อมูล

การระบวนการคำนวณค่าความคล้ายคลึงของคำศัพท์เป็นกระบวนการที่จะดำเนินการเมื่อทำการปรับโครงสร้างของฐานข้อมูลเป็นรูปแบบของคลาสและพร็อพเพอร์ตีทั้งฐานข้อมูลแล้ว จากนั้นข้อมูลแต่ละรายการจะถูกนำมาเข้ากระบวนการวัดค่าความคล้ายคลึง โดยมีชุดเมทาดาทาต้นแบบสำหรับการเปรียบเทียบที่ได้จากข้อมูลการประมวลผลการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอนและการศึกษาเมทาดาทาตามลักษณะของเลิร์นนิงอ็อบเจกต์เพิ่มเติม ตามกระบวนการในหัวข้อ 3.2.1.1 โดยกระบวนการเปรียบเทียบและผสานเมทาดาทาจะใช้แนวคิดการวัดความคล้ายคลึงด้วยแนวคิดของแจ็คการ์ด และการพิจารณาลักษณะข้อมูลที่จัดเก็บที่มีลักษณะที่เหมือนกัน โดยมีรายละเอียดดังต่อไปนี้

การคำนวณหาค่าความคล้ายคลึงระหว่างเอนทิตีหรือคลาสของแหล่งข้อมูลต้นทาง (Source) และเอนทิตีหรือคลาสเป้าหมาย (Target) โดยใช้การพิจารณาตามโครงสร้างความสัมพันธ์ของคำศัพท์ โดยการนำเครื่องมือการวัดความคล้ายคลึงตามโครงสร้างของคำจากเวิร์ดเน็ต (WordNet Similarity Package) ร่วมกับสมการการวัดความคล้ายคลึงของ Wu and Palmer (1994) โดยพิจารณาความลึกของ 2 กลุ่มคำ (Synset) ภายในโครงสร้างศัพท์สัมพันธ์ของเวิร์ดเน็ต ดังนี้

$$sim_{WuP}(c_1 c_2) = \frac{2*depth(LCS)}{(depth(c_1))+(depth(c_2))} \qquad (3\text{-}3)$$

จากสมการค่าของการคำนวณความคล้ายคลึงอยู่ระหว่าง 0 < ค่า Simwup <= 1 โดยค่าจากสมการจะมีค่ามากกว่าศูนย์เสมอ เนื่องจากค่าความลึกของ LCS ที่คำนวณได้ต้องไม่เท่ากับศูนย์ ทั้งนี้ค่าความคล้ายคลึงจะเท่ากับ 1 ได้นั้น เกิดจากกรณีที่คำทั้งสองอยู่ในซิงค์เซ็ตหรือกลุ่มคำศัพท์เดียวกัน

(3) การผสานเมทาดาทา

กระบวนการผสามาเมทาดาทาจะใช้การพิจารณาถึงผลคะแนนความคล้ายคลึงเพื่อประเมินว่าคลาสหรือพร็อพเพอร์ตีใดที่เหมือนกัน เกี่ยวข้องกัน หรือแตกต่างกัน จากนั้นจึงนำมาปรับปรุงเขตข้อมูลตามโครงสร้างความสัมพันธ์ของคลาสและพร็อพเพอร์ตี ดังนี้

(3.1) การพิจารณาและปรับปรุงเขตข้อมูลด้วยวิธีเชิงความหมาย

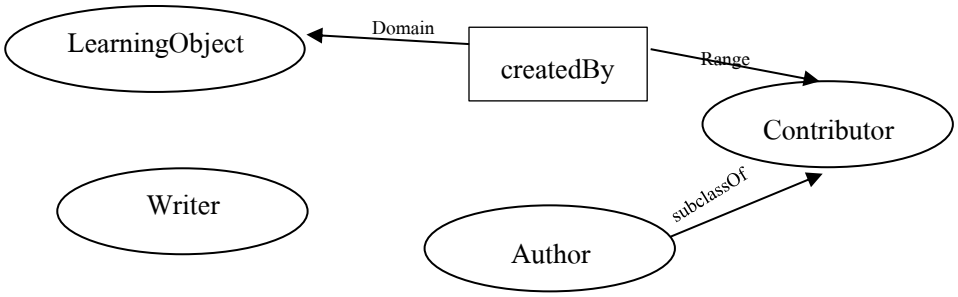
หลังจากการวัดความคล้ายคลึงของคลาสตามโครงสร้างคำศัพท์ด้วยเวิร์ดเน็ตแล้ว การกำหนดความสัมพันธ์ด้วยคุณสมบัติของอาว (OWL) เช่น owl:equivalentClass owl:equivalentProperty และ owl:sameAs เป็นต้น (Changqing, L., Ling, T. W., 2004; Seksun Suwanmanee, Djamal Benslimane and Philippe Thiran., 2005) เพื่อนำมาใช้ในการพิจารณาความสัมพันธ์ระหว่างคำศัพท์ที่ใช้ในเขตข้อมูล ดังรูปที่ 3.8

รูปที่ 3.8 การกำหนดคุณสมบัติระหว่างคำศัพท์ที่มีความคล้ายคลึงกัน

จากรูปที่ 3.8 การพิจารณาเขตข้อมูลภายในคลาสที่มีค่าความคล้ายคลึงกันด้วยคุณสมบัติ owl:sameAs เพื่อกำหนดว่าเขตข้อมูล Title และเขตข้อมูล xsd:name เหมือนกันหรือใช้แทนกันได้ ซึ่งวิธีการวิเคราะห์จะใช้วิธีการกำหนดชุดคำศัพท์ที่อยู่ในหมวดหมู่เดียวกันและสามารถใช้ร่วมกันหรือใช้แทนกันได้เป็นคลังคำ เพื่อนำมาวิเคราะห์ในการกำหนดความสัมพันธ์ข้างต้น ซึ่งชุดคลังคำนี้นำมาจากคำศัพท์ตามเค้าร่างเมทาดาทาของดับลินคอร์ และมาตรฐานลอมเป็นหลัก และเพิ่มคำศัพท์ตามการสำรวจการใช้งานของผู้ใช้จากตัวอย่างข้อมูลที่นำมาทดสอบ เพื่อให้การเปรียบเทียบเขตข้อมูลมีความถูกต้องและครอบคลุมคำศัพท์ทีเกี่ยวข้อง

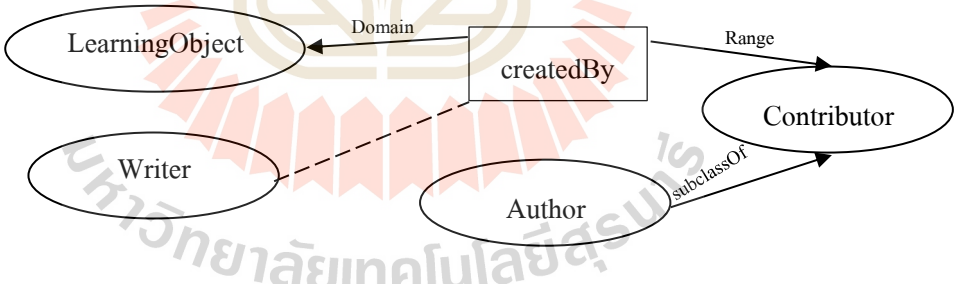
(3.2) การพิจารณารูปแบบความขัดแย้งเชิงโครงสร้างอ็อบเจกต์พร็อพเพอร์ตี

ในกระบวนการนี้ผู้วิจัยพิจารณาถึงโครงสร้างของคลาสทั้งสองโครงสร้างด้วยตนเองเพิ่มเติมจากการใช้การคำนวณความคล้ายคลึงเชิงความหมายของคำศัพท์จากศัพท์สัมพันธ์ของเวิร์ดเน็ต และการวิเคราะห์คำศัพท์ของเค้าร่างเมทาดาทา เพื่อกำหนดความสัมพันธ์ระหว่างเขตข้อมูลแต่ละเขตข้อมูล โดยกระบวนการนี้จะพิจารณาจากความสัมพันธ์ระหว่างคลาสหรืออ็อบเจกต์พร็อพเพอร์ตี (Object Properties) ที่เหมือนกันแต่ชื่อคลาสมีความแตกต่างกันอย่างสิ้นเชิง หรือมีค่าคะแนนความคล้ายคลึงต่ำกว่า 0.7 รวมถึงคลาสที่มีการจัดเก็บด้วยโครงสร้างของคลาสที่ซับซ้อน หรือคลาสที่มีโครงสร้างหลายระดับ ตามรูปที่ 3.9



รูปที่ 3.9 ลักษณะความขัดแย้งเชิงโครงสร้างสำหรับข้อมูลผู้แต่ง Writer และ Author

จากรูปที่ 3.9 แสดงให้เห็นถึงคลาสผู้แต่ง (Writer) ซึ่งบางครั้งสามารถใช้คำศัพท์ได้หลายคำ เช่น Author, Composer หรือเป็นกลุ่มผู้รับผิดชอบและเผยแพร่เนื้อหา (Contributor) และเมื่อวัดความคล้ายคลึงเชิงความหมายจากโครงสร้างศัพท์สัมพันธ์แล้วมีระยะทางที่ห่างกันมากทำให้คะแนนค่าความคล้ายคลึงต่ำ ก็จะใช้การพิจารณาอ็อบเจกต์พร็อพเพอร์ตีที่เหมือนกัน หรือสอดคล้องใกล้เคียงกันกัน และเพิ่มความสัมพันธ์ด้วยคุณสมบัติของภาษาอาว (OWL) อาทิ owl:equivalentClass owl:hasValue owl:someValuesFrom owl:allValuesFrom เข้ามาช่วยในการเชื่อมความสัมพันธ์ดังกล่าว ดังแสดงในรูปที่ 3.10



รูปที่ 3.10 ตัวอย่างการแก้ปัญหาความขัดแย้งเชิงโครงสร้าง

จากรูปที่ 3.10 สามารถนำคุณสมบัติ owl:allValueFrom ของภาษา OWL มาใช้ในการเชื่อมความสัมพันธ์ระหว่างคลาส writer เข้ากับคลาสอื่น ๆ ที่มีคุณสมบัติ createdBy เพื่อเป็นการจัดเก็บข้อมูลจากคลาส Contributor และ Author ร่วมกันได้ เนื่องจากทั้ง 2 คลาสนี้มีข้อมูลที่จัดเก็บคล้ายคลึงกันในลักษณะโครงสร้างการจัดเก็บข้อมูล

(4) การสร้างโครงสร้างชุดข้อมูลอาร์ดีเอฟ

เมื่อทำการคำนวณค่าความคล้ายคลึงและเลือกผสานเมทาดาทาตามกฎและเงื่อนไขการสร้างความสัมพันธ์ระหว่างคลาสและพร็อพเพอร์ตีแล้ว ผู้วิจัยจะทำการสร้างชุดข้อมูลอาร์อีเอฟจากรายละเอียดการแปลงและผสานเมทาดาทาจากกระบวนการข้างต้น โดยจะใช้เครื่องมือในการช่วยสร้างไฟล์อาร์ดีเอฟ ชื่อว่า D2RQ platform (D2RQ, 2012) โดยผู้วิจัยทำการสร้างไฟล์กำหนดเงื่อนไขการผสานคลาสและเมทาดาทาตามการออกแบบดังกล่าวข้างต้นไว้ก่อนล่วงหน้า เพื่อให้การแปลงฐานข้อมูลและการผสานเมทาดาทามีความถูกต้อง รวดเร็ว

```
map:contentobject__label a d2rq:PropertyBridge;
        d2rq:belongsToClassMap map:contentobject;
        d2rq:property rdfs:label;
        d2rq:pattern "contentobject #@@contentobject.CoId@@";
        .
map:contentobject_CoId a d2rq:PropertyBridge;
        d2rq:belongsToClassMap map:contentobject;
        d2rq:property vocab:contentobject_CoId;
        d2rq:propertyDefinitionLabel "contentobject CoId";
        d2rq:column "contentobject.CoId";
        .
map:contentobject_CoTitle a d2rq:PropertyBridge;
        d2rq:belongsToClassMap map:contentobject;
        d2rq:property vocab:contentobject_CoTitle;
```

รูปที่ 3.11 ตัวอย่างกฎการผสานเมทาดาทาและแปลงเป็นไฟล์อาร์ดีเอฟ

จากรูปที่ 3.11 เป็นชุดคำสั่งในการแปลงตารางข้อมูลจากฐานข้อมูลเชิงสัมพันธ์ และสามารถกำหนดไฟล์เงื่อนไขในการผสานเมทาดาทา ซึ่งในตัวอย่างเป็นการแปลงตาราง ให้อยู่ในรูปแบบคลาส และกำหนดให้แอตทริบิวต์ในตารางข้อมูลให้เป็นดาทาพร็อพเพอร์ตี (data properties) และกำหนดคำศัพท์ที่ต้องการเปรียบเทียบและใช้ในโครงสร้างอาร์ดีเอฟ ซึ่งใน

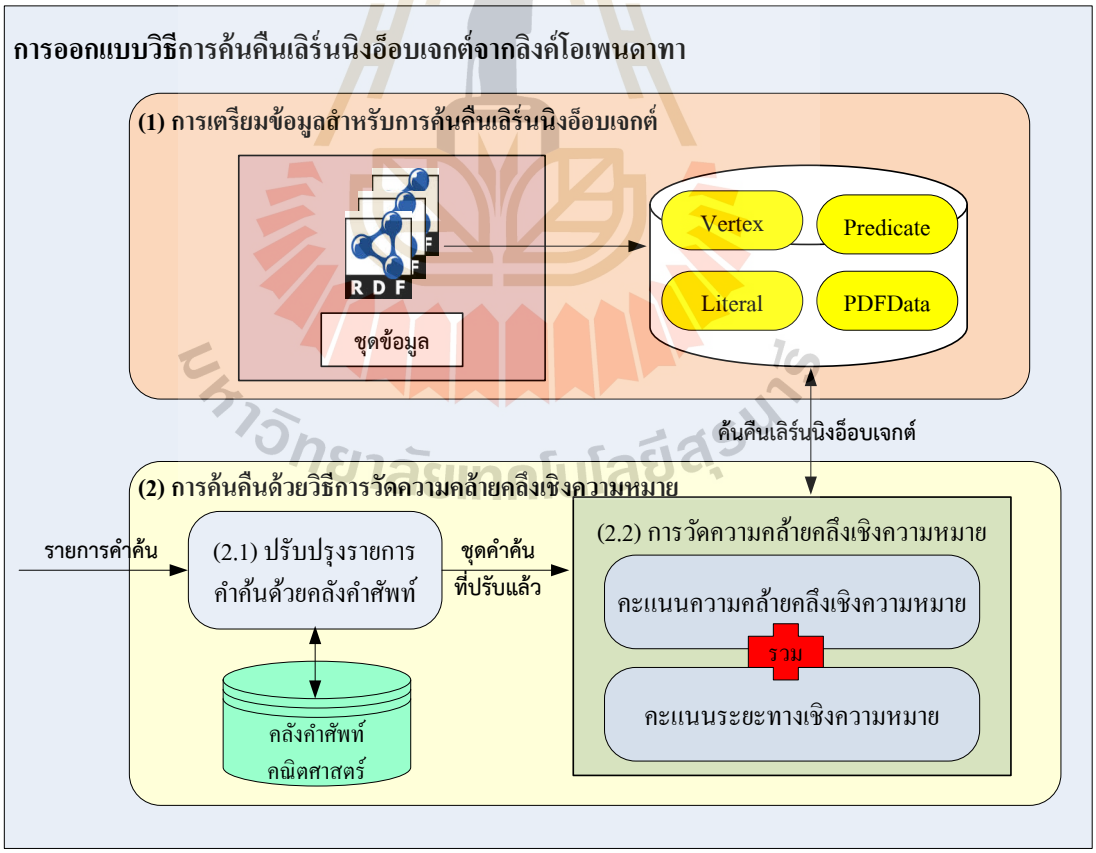
ตัวอย่างเป็นการใช้คำศัพท์ที่เครื่องมือกำหนดไว้เบื้องต้น คือ ชุดคำศัพท์ DC และ d2rq RDF ซึ่งเป็นคำศัพท์พื้นฐานในการสร้างคลาสและพร็อพเพอร์ตีในชุดข้อมูล

3.2.2 การออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์

ในกระบวนการค้นคืนเลิร์นนิงอ็อบเจกต์จะมีกระบวนการดังต่อไปนี้

3.2.2.1 การออกแบบวิธีการค้นคืนเลิร์นนิงอ็อบเจกต์

ขั้นตอนการค้นคืนเลิร์นนิงอ็อบเจกต์จะประกอบด้วยขั้นตอน 2 ขั้นตอนได้แก่ (1) การเตรียมข้อมูลสำหรับการค้นคืนเลิร์นนิงอ็อบเจกต์เพื่อช่วยเพิ่มประสิทธิภาพในการค้นคืนและลดการใช้ทรัพยากรของระบบ (2) การค้นคืนด้วยวิธีการวัดความคล้ายคลึงเชิงความหมาย โดยใช้แนวทางการคำนวณคะแนนความคล้ายคลึงเชิงความหมายของคำค้น และคำนวณคะแนนระยะทางเชิงความหมายของโหนดที่เกี่ยวข้องกับคำค้น เพื่อผลการสืบค้นที่ครอบคลุมเรื่องที่เกี่ยวข้อง โดยแสดงดังรูปที่ 3.12



รูปที่ 3.12 กรอบแนวคิดขั้นตอนการค้นคืนเลิร์นนิงอ็อบเจกต์

(1) การเตรียมข้อมูลสำหรับการค้นคืนเลิร์นนิงอ็อบเจกต์

การเตรียมข้อมูลก่อนการค้นคืน เมื่อฐานข้อมูลจากภายในระบบถูกแปลงให้เป็นชุดข้อมูลอาร์ดีเอฟในรูปแบบของรายการข้อมูลทริเปิล ก่อนที่ระบบจะพร้อมทำการค้นคืน คลังข้อมูลอาร์ดีเอฟจะถูกนำออกมาปรับโครงสร้างให้เป็นรูปแบบของตารางข้อมูลแทนรายการข้อมูลแบบทริเปิล เพื่อเพิ่มความรวดเร็วและลดทรัพยากรของระบบในการประมวลผลการค้นคืน โดยแนวคิดการเก็บข้อมูลในตารางแบบคีย์-แวลู (Key-Value) เพื่อกำหนดข้อมูลเป็นดัชนีในการชี้ข้อมูลในตาราง (Termehchy and Winslett, 2010) และยังช่วยให้การเพิ่มข้อมูลที่เข้ามาใหม่เป็นระบบต่อเนื่องไม่มีความซ้ำซ้อนหรือข้อมูลที่เก็บแบบกระจาย (Distribute) ที่อยู่คนละฐานข้อมูลก็สามารถจัดเก็บได้โดยไม่เกิดความซ้ำซ้อนระหว่างข้อมูลของแต่ละฐานข้อมูล ตารางข้อมูลที่สร้างขึ้นใหม่จะประกอบด้วยคอลเล็กชั่น 4 คอลเล็กชั่น ได้แก่ (1) คอลเล็กชั่นจุดยอด (Vertex Collection) สำหรับจัดเก็บประธาน (Subject) และค่าของคุณสมบัติ (Object) โดยชุดข้อมูลวัตถุสามารถเป็นประธานในความสัมพันธ์อื่นได้ จึงเก็บไว้ในตารางเดียวกัน (2) คอลเล็กชั่นภาคแสดง (Predicate Collection) เก็บข้อมูลภาคแสดงระหว่างประธานและข้อมูลแต่ละตัวเพื่อใช้เป็นข้อมูลเชื่อมโยงระหว่างโหนดต่าง ๆ (3) คอลเล็กชั่นข้อมูล (Literal Collection) สำหรับจัดเก็บข้อมูลที่ไม่ใช่คลาสที่เป็นค่าของคุณสมบัติ (Object) และ (4) ข้อมูลอาร์ดีเอฟ (RDF Data Collection) เก็บความสัมพันธ์รวมของชุดข้อมูลทั้งหมด เพื่อเชื่อมโยงรายละเอียดจากคอลเล็กชั่นทั้ง 3 เข้าด้วยกัน เปรียบเสมือนตารางบอกเส้นทางที่จะเป็นจุดเริ่มต้นในการค้นคืนข้อมูลในชุดข้อมูลอาร์ดีเอฟทั้งหมด โดยแสดงตัวอย่างข้อมูลทริเปิลตามตารางที่ 3.1

ตารางที่ 3.1 ตัวอย่างการเก็บข้อมูลแบบทริเปิลในตารางข้อมูลทั่วไป

Subject	Predicate	Object
Subject1	Predicate1	Object1
Subject1	Predicate2	Object2
Subject1	Predicate3	Object3

จากตารางที่ 3.1 เป็นข้อมูลแบบทริเปิลตัวอย่างที่เป็นข้อมูลจากชุดข้อมูลอาร์ดีเอฟ จะถูกนำไปแปลงไปเก็บในแบบคีย์-แวลู โดยแยกเป็นคอลเล็กชั่นแต่ละประเภท

ตารางที่ 3.2 การเก็บข้อมูลในแบบคีย์-แวลู คอลเล็กชั่นโหนดยอด (Vertex Collection)

VID	Value
0	Subject1
1	Object1
2	Object2
3	Object3

ตารางที่ 3.3 การเก็บข้อมูลในแบบคีย์-แวลู คอลเล็กชั่นพร็อพเพอร์ตี (Predicate Collection)

PID	Value
0	Predicate1
1	Predicate2
2	Predicate3

ตารางที่ 3.4 การเก็บข้อมูลอาร์ดีเอฟ (RDFData)

VID	Incoming Node	Outgoing Node
[0]	{}	[1,2,3]
[1]	{“0”: [0]}	[ ]
[2]	{“1”: [0]}	[ ]
[3]	{“2”: [0]}	[ ]

จากตารางที่ 3.2 และ 3.3 เป็นการบันทึกข้อมูลจากข้อมูลแบบทริเปิลให้เป็นคีย์-แวลู เพื่อสร้างดัชนีเข้าถึงข้อมูลทั้งหมดผ่านโหนดที่แทนด้วยตัวเลข และสร้างตารางที่ 3.4 เป็นตารางรวมความสัมพันธ์จากชุดข้อมูลอาร์ดีเอฟโดยนำข้อมูลจากคอลเล็กชั่นจุดยอด และคอลเล็กชั่นพร็อพเพอร์ตีที่เชื่อมโยงกันมาสร้างเป็นตารางข้อมูลชุดใหม่เพื่อความสะดวกในการเข้าถึงข้อมูลแต่ละโหลดผ่านเลขดัชนีแทนการเข้าถึงข้อมูลตามลำดับแถวแบบปกติ และรูปแบบตารางนี้สามารถรองรับข้อมูลเอิร์นนิงอ็อบเจกต์รายการใหม่ได้โดยไม่กระทบกับข้อมูลเดิม โดยข้อมูลใหม่ถูกแปลงและบันทึกลงตารางเหล่านี้ต่อไปตามลำดับ

(2) การค้นคืนด้วยวิธีการวัดความคล้ายคลึงเชิงความหมาย

เทคนิควิธีในการค้นคืนเลิร์นนิงอ็อบเจกต์ประกอบด้วยการปรับปรุงคำค้นจากรายการคำค้นที่ใส่เข้ามา จากนั้นจะนำไปคำนวณคะแนนความคล้ายคลึงเชิงความหมายและการวัดระยะทางเชิงความหมายจากตารางข้อมูลในกระบวนการเตรียมข้อมูลการค้นคืน

(2.1) การปรับปรุงคำค้นด้วยคลังคำศัพท์คณิตศาสตร์

เมื่อทำการเตรียมข้อมูลของเลิร์นนิงอ็อบเจกต์ในรูปแบบตารางที่กำหนดแล้ว จะสามารถทำการค้นคืนได้ โดยเมื่อผู้ใช้ใส่คำค้นเข้าสู่ระบบและทำการปรับปรุงคำค้นโดยคลังคำศัพท์ทางคณิตศาสตร์ ระบบจะทำการเปรียบเทียบคำค้นกับคลังคำ เพื่อหาคำค้นที่เหมาะสมและหัวเรื่องที่เกี่ยวข้อง การจัดทำคลังคำศัพท์ที่ใช้เป็นตัวอย่างทดลองการวิจัยครั้งนี้ ใช้การจัดเก็บคำศัพท์เชิงสัมพันธ์ด้วยกรอบแนวคิดสคอส (Simple Knowledge Organization System: SKOS) โดยโครงสร้างของคำศัพท์ในสคอสมีการกำหนดความสัมพันธ์ของกลุ่มคำศัพท์ใน 3 กลุ่ม ได้แก่ คำศัพท์ที่มีความสัมพันธ์ในลักษณะเท่าเทียมหรือเป็นสิ่งเดียวกัน (Alternative Term) คำศัพท์ที่มีความสัมพันธ์ในลักษณะลดหลั่นในระดับชั้นที่สูงกว่าและต่ำกว่า (Broader Term, Narrower Term) และความสัมพันธ์ในลักษณะที่เกี่ยวข้องกันหรือสัมพันธ์กันในบางประเด็น (Related Term) ซึ่งคลังคำศัพท์ที่นำมาใช้เป็นการรวบรวมขึ้นใหม่จากแหล่งข้อมูลหลายแหล่งซึ่งเป็นชุดคำศัพท์ที่เกี่ยวข้องกับรายวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 1 ถึง 6 รวมทั้งสิ้น 380 คำ โดยแหล่งรวบรวมคำศัพท์ประกอบด้วย 1) สำกัดจากหนังสือเรียนรายวิชาพื้นฐานคณิตศาสตร์ และหนังสือเรียนรายวิชาเพิ่มเติมคณิตศาสตร์ ระดับชั้นมัธยมศึกษาชั้นปีที่ 1 ถึง 6 ตามการกำหนดของหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน จำนวน 16 เล่ม 2) พจนานุกรมคำศัพท์คณิตศาสตร์สำหรับนักเรียน 3) คำศัพท์สัมพันธ์จากเวิร์ดเน็ตจากคำหลักคณิตศาสตร์และคำที่เกี่ยวข้อง และ 4) คำศัพท์ที่สกัดจากเลิร์นนิงอ็อบเจกต์ตัวอย่างที่ใช้ในการทดลองมอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์จำนวน 300 รายการ โดยคลังคำศัพท์ที่รวบรวมขึ้นนี้ผ่านการประเมินจากผู้เชี่ยวชาญจำนวน 5 คน และครูผู้สอนรายวิชาคณิตศาสตร์ระดับมัธยมศึกษารวมจำนวน 5 ท่าน ทั้งนี้ ชุดคำศัพท์ทั้งหมดจะนำมากำหนดเป็นโครงสร้างคำศัพท์สัมพันธ์ โดยมีการกำหนดโครงสร้างด้วยแนวคิดสคอส (SKOS) และจัดเรียงคำศัพท์ตามความสัมพันธ์ภายในหัวเรื่องคณิตศาสตร์จากศัพท์สัมพันธ์เวิร์ดเน็ต

ชุดคำศัพท์สัมพันธ์ที่จัดทำขึ้นนี้จะเป็นชุดคำศัพท์ที่จะนำไปใช้ในขั้นตอนการพัฒนามอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์ โดยการนำไปประกอบการพิจารณาเปรียบเทียบคำค้น เพื่อช่วยในการเลือกคำค้นที่เหมาะสม และกำหนดหัวเรื่องที่เกี่ยวข้องกับคำค้นดังกล่าวเพื่อสืบค้นหัวเรื่องที่เกี่ยวข้อง นอกจากนี้ยังมีส่วนในการวัดความคล้ายคลึงเชิงความหมายของผลการสืบค้นโดยการใช้คำค้นเปรียบเทียบกับดัชนีตัวแทนเอกสารแต่ละรายการที่ค้นได้ เพื่อนำไปใช้ในการจัดลำดับการนำเสนอและจัดกลุ่มเนื้อหาตามเมทาดาทาที่ได้จากการสำรวจความต้องการของผู้ใช้

(2.2) การวัดความคล้ายคลึงเชิงความหมาย

การสืบค้นด้วยคำค้นทั่วไปในชุดข้อมูลโอเพนดาทาจะมีปัญหาเรื่องความแม่นยำต่ำ เนื่องจากลักษณะข้อมูลที่เก็บในชุดข้อมูลสามารถเชื่อมโยงไปยังชุดข้อมูลอื่น ๆ ได้ ทำให้ขอบเขตการสืบค้นด้วยคำค้นเพียงคำเดียวสามารถดึงข้อมูลที่เกี่ยวข้องได้ในจำนวนที่มากเกินจริง ดังนั้นจึงมีงานวิจัยหลายชิ้นที่คิดค้นวิธีกระบวนการสืบค้นโดยการกำหนดขอบเขตหรือระดับความลึก (K-Depth) ในการสืบค้นข้อมูล ซึ่งจะเกิดปัญหาของการกำหนดระดับความลึกในการสืบค้นที่ไม่เหมาะสม ได้แก่การกำหนดขอบเขตการค้นที่แคบเกินไป การเปรียบเทียบคำค้นมีความแม่นยำมาก แต่ผลลัพธ์ที่เกี่ยวข้องจะไม่ถูกสืบค้นขึ้นมาได้ครบถ้วน หรือการกำหนดความลึกมากเกินไปก็จะส่งผลต่อจำนวนผลลัพธ์ที่ออกมามากเกินความจำเป็น และส่งผลต่อประสิทธิภาพของระบบอย่างมาก เป็นต้น ดังนั้นในกระบวนการนี้จะเป็นการใช้การวัดความคล้ายคลึงเชิงความหมายช่วยในการกำหนดระดับความลึกในการสืบค้นภายในโครงสร้างอาร์ดีเอฟที่เหมาะสมที่สุด

กระบวนการวัดความคล้ายคลึงเชิงความหมายจะทำการคำนวณค่าคะแนนความคล้ายคลึงจาก 2 ส่วน ได้แก่

(2.2.1) การวิเคราะห์ความสัมพันธ์ของคุณสมบัติของทรัพยากร (Relational Predicate Keyword) จากรูปที่ 3.13 คลาสแต่ละคลาสจะสัมพันธ์กันด้วยคุณสมบัติของทรัพยากร โดยจะมีคำศัพท์ที่ใช้บ่งชี้ความสัมพันธ์ เช่น ID, Title, Level เป็นต้น ซึ่งคำศัพท์เหล่านี้สามารถบ่งบอกถึงลักษณะความสัมพันธ์ของตัวทรัพยากร (Subject) และค่าของคุณสมบัติ (Object) แต่ละตัวได้ และบ่งบอกถึงขอบเขตและเนื้อหาของกลุ่มคำศัพท์ในเชิงความหมายได้ เช่น ถ้าเป็นกลุ่มของ Friend of A Friend (FOAF) จะเกี่ยวข้องกับการบรรยายคุณสมบัติของบุคคล หรือ Semantically Inter-Linked Online Community (SICO) จะเป็นกลุ่มคำศัพท์ที่ใช้อธิบายเกี่ยวกับข้อมูลบนบล็อก (Blog) หรือกระดานข่าว (Webboard) หรือกลุ่มสังคมออนไลน์บนเว็บไซต์ต่าง ๆ ดังนั้นงานวิจัยนี้จึงนำคำศัพท์ของคุณสมบัติทรัพยากรนี้มาใช้ในการวิเคราะห์ความหมายของคำเพื่อใช้ในการวัดความคล้ายคลึง ซึ่งการวัดความคล้ายคลึงนี้จะได้แนวคิดการวัดความคล้ายคลึงของลิน (Lin, 1998) เนื่องจากเป็นการวัดความคล้ายคลึงของข้อมูลในโครงสร้างที่มีความสัมพันธ์ระหว่างคำศัพท์

$$sim_{Lin} = \frac{2.IC(LCS)}{IC(concept_1)+IC(concept_2)} \qquad (3\text{-}4)$$

โดย

IC คือ ค่า Information Concept โดยคิดจากความน่าจะเป็นที่ปรากฏ concept ทั้งสอง

LCS คือ คำที่เชื่อมกันใกล้สุดระหว่าง concept ที่ 1 และ 2

กระบวนการในการวัดความคล้ายคลึง เริ่มจากผู้ใช้ใส่คำสำคัญในการค้นเข้ามา จะนำคำค้นนั้นไปเปรียบเทียบกับโหนดที่เป็นค่าของคุณสมบัติของทรัพยากรแบบอักขระ (Literal) การเปรียบเทียบใช้การเปรียบเทียบแบบจับคู่ (Matching) เพื่อให้ได้ค่าที่ตรงกับคำค้นมากที่สุด และเก็บค่า Vertex ID ของค่าของคุณสมบัติประเภทอักขระ (Literal) นั้นเพื่อเชื่อมโยงกลับไปที่โหนดหลักผ่าน PID ที่เป็นรหัสของดาทาพร็อพเพอร์ตี (Data Properties) เมื่อเชื่อมข้อมูลสำเร็จก็จะได้ชุดของคำศัพท์คุณสมบัติ (List of Predicate) ที่จะนำไปใช้ในการวัดระยะทางของคำศัพท์ได้

เมื่อได้ชุดของคำศัพท์แล้วจะนำไปคำนวณหาคะแนนระยะทาง (Distance Score) โดยกำหนดให้จุดโหนดที่เป็นจุดเริ่มต้นมีค่าเท่ากับ 0 และเดินทางไปยังโหนดแต่ละระดับมีค่าเท่ากับ 1 ในแต่ละชั้นของโครงสร้าง โดยใช้สมการการหาคะแนนระยะทาง (3-5) โดยประยุกต์จากสูตรของแบ (Bae et al., 2013)

$$SDS(k, o_i) = 1 - \frac{NSC(k,o_i)-1}{Max[NSC(k,o_1)NSC(k,o_n)]} \qquad (3\text{-}5)$$

เมื่อ

SDS คือ คะแนนระยะทางระหว่างคำค้นกับโหนด

k คือ คำค้นที่ใส่เข้าระบบ

o_i คือ ชุดข้อมูลที่เชื่อมโยงกันในอาร์ดีเอฟ

NSC คือ ค่าของการเปลี่ยนตำแหน่งของโหนดข้อมูล

(2.2.2) คะแนนการเปรียบเทียบเชิงความหมายระหว่างคลาส (Semantic Similarity Score) นอกจากการวัดระยะทางเชิงความหมายจากกระบวนการที่ 1 แล้ว ในงานวิจัยนี้ยังได้ใช้การเปรียบเทียบความคล้ายคลึงเชิงความหมายเพื่อนำไปคำนวณร่วมกับการวัดระยะทางเชิงความหมายเพื่อใช้ในการประเมินความเหมาะสมของระดับการสืบค้นทั้งรายละเอียดของคลาสกับคำค้น และระยะความลึกของการสืบค้นในโครงสร้างการจัดเก็บข้อมูลอีกด้วย ดังนั้นในส่วนการเปรียบเทียบเชิงความหมายของคำค้นกับคลาสแต่ละคลาสจะใช้การประยุกต์จากสูตรของลิน (Lin, 1998) ตามสมการที่ 3-5 และการเปลี่ยนตัวแปรจากการเปรียบเทียบคำค้นต่อคำในเอกสารให้เป็นคำค้นและอัตราส่วนของเขตข้อมูลในคลาสแต่ละคลาส ตามแนวคิดของแบ (Bae, Nguyen, Kang and Oh, 2013)

$$SSS(k, o_i) = \frac{2 \times \log r(c_o)}{\log r(c_i) + \log r(c_k)} \qquad (3\text{-}6)$$

เมื่อ

k คือ คำค้น

c_k คือ จำนวนของคลาสที่มีคำค้นปรากฏ

c_i คือ จำนวนคลาสที่มีอ็อบเจกต์ปรากฏ

c_o คือ จำนวนคลาสที่มีปรากฏคำค้นและอ็อบเจกต์

จากสมการที่ 3-6 เป็นการเปรียบเทียบความคล้ายของคำค้นกับคลาสที่มีค่าของคุณสมบัติประเภทอักขระ (Literal) ปรากฏอยู่ โดย $r(c_k)$ คืออัตราส่วนระหว่างคลาสที่มีคำค้นปรากฏอยู่กับจำนวนคลาสทั้งหมด $r(c_i)$ คืออัตราส่วนระหว่างคลาสที่มีค่าของคุณสมบัติของทรัพยากรที่ต้องการเปรียบเทียบปรากฏอยู่ กับจำนวนคลาสทั้งหมด และ $r(c_o)$ เป็นอัตราส่วนระหว่างจำนวนคลาสที่ปรากฏทั้งคำค้นและค่าของคุณสมบัติของทรัพยากรในคลาสนั้น กับจำนวนคลาสทั้งหมด โดยการเปรียบเทียบความคล้ายคลึงเชิงความหมายนี้จะทำการเปรียบเทียบจากคำค้นและโครงสร้างของคลังคำศัพท์ที่สร้างไว้ในขั้นตอนการสร้างคลังศัพท์สัมพันธ์

จากการวัดความคล้ายคลึงทั้งทางด้านระยะทางเชิงความหมาย (Semantic Distance Score: SDS) และการวัดความคล้ายคลึงเชิงความหมายระหว่างคลาส (Semantic Similarity Score: SSS) ทางผู้วิจัยจึงนำการวัดความคล้ายคลึงทั้งสองส่วนมาใช้ร่วมกัน โดยมีค่าพารามิเตอร์ (p) เพื่อเป็นสัมประสิทธิ์ในการกำหนดค่าผกผันระหว่างการวัดระยะทางเชิงความหมายและการวัดความคล้ายคลึงเชิงความหมายให้มีค่าอยู่ระหว่าง 0 ถึง 1 โดยการกำหนดค่าพารามิเตอร์ p ที่เหมาะสมเพื่อให้ประสิทธิภาพการสืบค้นมีค่าสูงที่สุด โดยการทดสอบปรับเปลี่ยนค่า p เพื่อให้ได้ค่าที่เหมาะสมกับชุดข้อมูลอาร์ดีเอฟแต่ละชุดที่แตกต่างกัน

$$SS(k, o_i) = p \cdot SDS(k, o_i) + (1 - p) \cdot SSS(k, o_i) \qquad (3\text{-}7)$$

เมื่อ

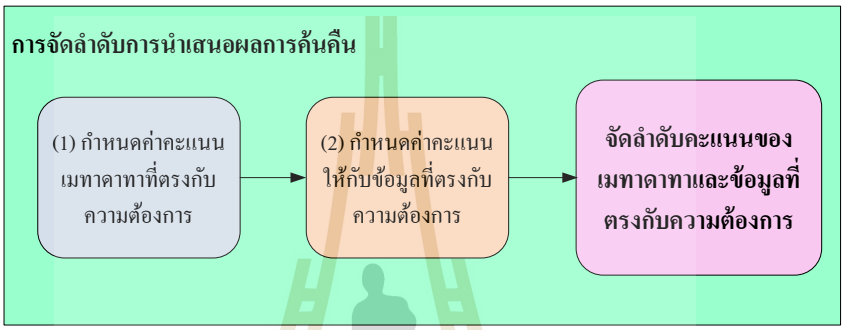
k คือ คำค้น

o_i คือ อ็อบเจกต์ที่ทำการเปรียบเทียบ

p คือ ค่าพารามิเตอร์ให้น้ำหนัก

3.2.2.2 การจัดลำดับการนำเสนอเลิร์นนิงอ็อบเจกต์ (Ranking)

หลังการคำนวณความคล้ายคลึงเชิงความหมายโดยมีส่วนประกอบทางด้านระยะทางเชิงความหมายและความคล้ายคลึงของคลาสเชิงความหมายแล้ว งานวิจัยนี้ยังได้เพิ่มการจัดการรายการผลการสืบค้นในส่วนการนำเสนอให้มีความสอดคล้องกับความต้องการของผู้ใช้ โดยวิเคราะห์เค้าร่างเมทาดาทาและข้อมูลตัวแทนเอกสารทางด้านสถิติของระเบียน เพื่อนำมาใช้ในการนำเสนอผลการสืบค้นได้ดียิ่งขึ้น ดังรูปที่ 3.13



รูปที่ 3.13 กรอบแนวคิดการจัดลำดับการนำเสนอผลการสืบค้น

การจัดลำดับการนำเสนอในงานวิจัยนี้ประกอบด้วยแนวทางการจัดลำดับผลการสืบค้นโดยวิธีการ Query Dependent Ranking (Liu, 2009) โดยการจัดลำดับผลการสืบค้นโดยนำองค์ประกอบเมทาดาทา และข้อมูลทางบรรณานุกรมของเลิร์นนิงอ็อบเจกต์แต่ละรายการมาใช้กำหนดคะแนน โดยการเปรียบเทียบเมทาดาทาที่ต้องการร่วมกับข้อมูล (Vaule) ของรูปแบบเลิร์นนิงอ็อบเจกต์ (LOFormat) หรือประเภทของเลิร์นนิงอ็อบเจกต์ (LOType) และใช้กระบวนการเพิ่มค่าน้ำหนักให้กับองค์ประกอบเมทาดาทาและข้อมูลทางบรรณานุกรม (Bibliographic) ตามที่ผู้ใช้ต้องการให้มีค่าคะแนนสูงขึ้น เพื่อนำไปเปรียบเทียบคะแนนเพื่อการเรียงลำดับผลการสืบค้นที่มีรายละเอียดของรูปแบบและประเภทเลิร์นนิงอ็อบเจกต์ตรงกับความต้องการของผู้ใช้

กระบวนการเรียงลำดับจะใช้สมการคำนวณหาความคล้ายคลึงระหว่างข้อมูลสองรายการแบบปริภูมิเวกเตอร์ (Vector Space) ด้วยวิธีการวัดความคล้ายคลึงแบบโคซายน์ (Cosine Similarity) จากนั้นจะเพิ่มค่าน้ำหนักตามเงื่อนไขการเรียงลำดับของผู้ใช้ เพื่อให้เกิดค่าคะแนนที่แตกต่างกันของผลการสืบค้น และสามารถจัดลำดับตามค่าน้ำหนักของข้อมูลบรรณานุกรมที่ตรงความต้องการของผู้ใช้แต่ละเงื่อนไข (Zhou, Lawless, Min and Wade, 2010) โดยใช้สมการดังนี้

$$sim(d,q) = \frac{[\sum_{k=1}^{n} d*q]}{\sqrt{\sum_{k=1}^{n} d^2} * \sqrt{\sum_{k=1}^{n} q^2}} * F \qquad (3\text{-}8)$$

เมื่อ

d คือ ค่าน้ำหนักของโหนดที่พบ

q คือ ค่าน้ำหนักของคำสำคัญที่ใช้ค้นหา

F คือ ค่าพารามิเตอร์ให้น้ำหนักเมทาดาทาที่ต้องการ

จากสมการที่ 3-8 เป็นสมการหาความคล้ายคลึงของคำสองคำ โดยเปรียบเทียบความคล้ายเชิงเส้นคือไม่ยึดความหมายของคำแต่วัดความคล้ายคลึง โดยการให้ค่าน้ำหนักของโหนดที่พบในโครงสร้างอาร์ดีเอฟ และการให้น้ำหนักคำสำคัญที่ใช้สืบค้น นอกจากนี้ยังมีการให้ค่าน้ำหนักเมทาดาทาที่ต้องการสำหรับเรียงลำดับโดยการคูณด้วยค่าคงที่ ระหว่าง 0.1 ถึง 1.0 โดยเลือกให้น้ำหนักกับเมทาดาทาที่ต้องการ โดยผลการสำรวจความต้องการของผู้ใช้ พบว่าเขตข้อมูลที่ผู้ใช้นิยมใช้เพื่อพิจารณาเลือกผลการสืบค้นหรือจัดกลุ่มเนื้อหาการสืบค้น ได้แก่ 1) ชนิดของไฟล์เอกสาร เช่น เอกสารเวิร์ด รูปภาพ ไฟล์เสียง เป็นต้น 2) ภาษาที่ใช้ในตัวสื่อ 3) ระดับชั้นปีการศึกษา 4) รูปแบบเอกสาร เช่น เอกสารฉบับเต็ม มีไฟล์แนบหรือไม่ 5) หัวเรื่องคณิตศาสตร์ที่อยู่ในเนื้อหา เป็นต้น ซึ่งข้อมูลเหล่านี้ช่วยให้ผู้ใช้ตัดสินใจได้ว่าจะเข้าไปดูรายละเอียดรายการสืบค้นนั้นหรือไม่ เพื่อลดเวลาการพิจารณา ผู้วิจัยจึงนำปัจจัยเหล่านี้มาเป็นปัจจัยที่ใช้ในการจัดลำดับเนื้อหา เพื่อให้ผู้ใช้พบรายการที่ตนเองต้องการรวดเร็วมากยิ่งขึ้น

3.3 การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

การประเมินโมเดล โดยการพัฒนาระบบต้นแบบการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ตามโมเดลที่พัฒนาขึ้น จากนั้นทำการประเมิน 2 วิธี คือ (1) การประเมินประสิทธิภาพของระบบโดยการทดสอบ และ (2) การประเมินความสามารถในการทำงานของระบบโดยผู้เชี่ยวชาญ ดังนี้

3.3.1 การประเมินประสิทธิภาพของระบบโดยการทดสอบ

การประเมินประสิทธิภาพของระบบทำโดยการประเมินความถูกต้องของระบบการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ ใช้วิธีการประเมินความถูกต้องของระบบ 2 วิธี ดังนี้

1) การวัดค่ามาตรฐานในการประเมินความถูกต้อง 3 ค่า ได้แก่ ค่าความแม่นยา (Precision) ค่าการจำได้ (Recall) และอัตราการรู้จา (F-measure) มีสูตรคำนวณค่าต่าง ๆ ดังสมการที่ 3-9 ถึง 3-11

2) การเปรียบเทียบความสามารถกับระบบการจัดการข้อมูลอาร์ดีเอฟเจนาเฟรมเวิร์ค (Jena Framework)

Precision คือค่าความแม่นยำ ที่จะแสดงให้เห็นว่าระบบที่พัฒนาขึ้นมามีความแม่นยำเพียงใด โดยสามารถคำนวณได้จากสูตรดังสมการ

$$Precision = \frac{TP}{(TP+FP)} \qquad (3\text{-}9)$$

เมื่อ

ค่า *TruePositive* หรือ *TP* หมายถึง เอกสารที่เกี่ยวข้องและค้นคืนได้

ค่า *False Negative* หรือ *FN* หมายถึง เอกสารที่เกี่ยวข้องแต่ค้นคืนไม่ได้

Recall คือค่าการจำได้ ที่แสดงให้เห็นว่าเมื่อระบบได้ทำการดึงคำตอบออกมาแล้วมีความถูกต้องเพียงใด โดยสามารถคำนวณได้จากสูตรดังสมการ

$$Recall = \frac{TP}{(TP+FN)} \qquad (3\text{-}10)$$

เมื่อ

ค่า *TruePositive* หรือ *TP* หมายถึง เอกสารที่เกี่ยวข้องและค้นคืนได้

ค่า *False Negative* หรือ *FN* หมายถึง เอกสารที่เกี่ยวข้องแต่ค้นคืนไม่ได้

F-measure คือ อัตราการรู้จำ ซึ่งเป็นค่าเฉลี่ยที่ให้ความสำคัญกับความแม่นยำและค่าการจำได้ คำนวณได้จากสูตรดังสมการ

$$F-Measure = \frac{(2*Precision*Recall)}{(Precision+Recall)} \qquad (3\text{-}11)$$

การเปรียบเทียบความสามารถในการทำงานของระบบกับระบบจัดการข้อมูลอาร์ดีเอฟเจนาเฟรมเวิร์ค โดยการเปรียบเทียบจะเปรียบเทียบการทำงานทั้งสองระบบ โดยใช้การเปรียบเทียบเวลาในการทำงาน (Time Response) โดยใช้กลุ่มข้อมูลตัวอย่างที่มีขนาดแตกต่างกัน ตั้งแต่ 1MB 10MB 30MB 50MB 70MB และ 100MB โดยบันทึกเวลาการตอบสนองของทั้งสองระบบและเปรียบเทียบผลในรูปแบบกราฟ

มหาวิทยาลัยเทคโนโลยีสุรนารี

3.3.2 การประเมินความสามารถในการทำงานของระบบโดยผู้เชี่ยวชาญ

1) วิธีการประเมิน การประเมินความสามารถของระบบจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ในขั้นตอนนี้ ทำโดยให้ผู้เชี่ยวชาญทดลองใช้ระบบต้นแบบที่ติดตั้งลงบนเครื่องคอมพิวเตอร์ โดยผู้วิจัยกำหนดสถานการณ์ในการค้นคืนข้อมูลเป็นโจทย์ให้ผู้ใช้ทดลองใช้งาน จากนั้น ผู้วิจัยจะนำแบบประเมินที่ผู้วิจัยพัฒนาขึ้นให้ผู้เชี่ยวชาญทำการประเมินความสามารถของระบบ

2) กลุ่มตัวอย่าง ได้แก่ ผู้เชี่ยวชาญด้าน คอมพิวเตอร์จำนวน 5 คน (Neilsen, 2012) และครูผู้สอนรายวิชาคณิตศาสตร์ในหลักสูตรสองภาษา จำนวน 15 คน

3) เครื่องมือที่ใช้ในการรวบรวมข้อมูล ได้แก่ แบบประเมินความสามารถของระบบ ที่ผู้วิจัยพัฒนาขึ้น ประกอบด้วยข้อคำถาม 2 ตอน ดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ลักษณะแบบสอบถามเป็นแบบตรวจสอบรายการ (Checklist) จำนวน 2 ข้อ ประกอบด้วย เพศ และ ระดับการศึกษา

ตอนที่ 2 ความคิดเห็นเกี่ยวกับความสามารถของระบบ ลักษณะแบบสอบถามเป็นแบบมาตราส่วนประมาณค่า (Rating Scale) จำนวน 21 ข้อ โดยมีเกณฑ์การให้คะแนน ตามวิธีการของ ลิเคอร์ท (Likert) มี 5 ระดับ (บุญชม ศรีสะอาด, 2545: 102-103) ดังนี้

5 หมายถึง เหมาะสมมากที่สุด
4 หมายถึง เหมาะสมมาก
3 หมายถึง เหมาะสมปานกลาง
2 หมายถึง เหมาะสมน้อย
1 หมายถึง เหมาะสมน้อยที่สุด

โดยกำหนดรูปแบบการแปลความหมายของช่วงคะแนนได้ดังนี้

4.50- 5.00 หมายถึง ความต้องการในการสืบค้นข้อมูลอยู่ในระดับมากที่สุด
3.50 - 4.49 หมายถึง ความต้องการในการสืบค้นข้อมูลอยู่ในระดับมาก
2.50 - 3.49 หมายถึง ความต้องการในการสืบค้นข้อมูลอยู่ในระดับปานกลาง
1.50 - 2.49 หมายถึง ความต้องการในการสืบค้นข้อมูลอยู่ในระดับน้อย
1.00 - 1.49 หมายถึง ความต้องการในการสืบค้นข้อมูลอยู่ในระดับน้อยที่สุด

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตอนที่ 3 ข้อมูลด้านข้อเสนอแนะอื่น ๆ โดยใช้คำถามปลายเปิด (open-ended questions) เพื่อให้ผู้ตอบแบบสอบถามแสดงความคิดเห็นเพิ่มเติม

4) การทดสอบประมิทธิภาพของเครื่องมือ นำแบบประเมินที่สร้างขึ้นเสนอต่อผู้เชี่ยวชาญ 3 ท่าน เพื่อตรวจสอบความสอดคล้องระหว่างข้อคาถามกับจุดประสงค์ของการประเมิน โดยประเมินค่าดัชนีความสอดคล้อง (Item-Objective Congruence Index: IOC) ค่าความสอดคล้องระหว่างข้อคำถามของแบบประเมินกับจุดประสงค์ในการประเมินความสามารถของระบบ ผลการคำนวณค่า IOC อยู่ระหว่าง 0.8-1.0 ซึ่งผ่านเกณฑ์ที่กำหนดให้ข้อคาถามที่มีค่าคะแนน IOC ตั้งแต่ 0.5 ขึ้นไปสามารถใช้เป็นข้อคำถามได้

5) การวิเคราะห์ข้อมูล ทำการวิเคราะห์ข้อมูลด้วยการวิเคราะห์ค่าความถี่ (Frequency) ความเบี่ยงเบนมาตรฐาน (Standard Deviation) และหาค่าเฉลี่ย (Mean)

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 4
ผลการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตของผู้ใช้

งานวิจัยนี้มุ่งเน้นพัฒนาโมเดลสำหรับจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา เพื่อสร้างเป็นแหล่งรวบรวมเลิร์นนิงอ็อบเจกต์ไว้แหล่งเดียวโดยผู้ใช้สามารถจัดเก็บเลิร์นนิงอ็อบเจกต์และค้นคืนได้ตรงความต้องการมากที่สุด ดังนั้น เพื่อให้ระบบสามารถตอบสนองพฤติกรรมการใช้งานของผู้ใช้ ผู้วิจัยจึงทำการศึกษาพฤติกรรมการค้นคืนและใช้งาน เลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตของผู้ใช้ ซึ่งได้แก่ ครูผู้สอนรายวิชาคณิตศาสตร์ระดับมัธยมศึกษา ในสังกัดกระทรวงศึกษาธิการ โดยเป็นโรงเรียนที่เข้าโครงการการจัดการเรียนการสอนเป็นภาษาอังกฤษตามที่กระทรวงศึกษาธิการประกาศไว้ ประกอบไปด้วยโรงเรียนสอนหลักสูตร สองภาษาในหลักสูตรอิงลิชโปรแกรม (English Program) และมินิอิงลิชโปรแกรม (Mini English Program) โรงเรียนมาตรฐานสากล (World Class Standard School) ทั้งนี้ เพื่อเป็นข้อมูลเบื้องต้นในการออกแบบและพัฒนาโมเดลต่อไป

การศึกษาใช้วิธีการวิจัยเชิงสำรวจ (Survey Research) เก็บรวบรวมข้อมูลโดยใช้แบบสอบถาม กลุ่มประชากรได้แก่ ครูอาจารย์ที่รับผิดชอบหลักสูตรการเรียนการสอนสองภาษา (English Program) จากโรงเรียนในกำกับดูแลของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ทั่วประเทศ จำนวน 100 แห่ง และโรงเรียนเอกชน จำนวน 50 แห่ง รวมทั้งสิ้น 150 แห่ง โดยเป็นโรงเรียนที่มีการเรียนการสอนในระดับมัธยมศึกษาตอนต้นและมัธยมศึกษาตอนปลายในหลักสูตรสองภาษา (English Program) เนื่องจากจำนวนโรงเรียนทั้งหมดมีเพียง 150 โรงเรียน ดังนั้นจึงกำหนดกลุ่มตัวอย่างโดยการสุ่มเลือกครูคณิตศาสตร์ที่สอนในหลักสูตรสองภาษามาโรงเรียนละ 2 คน ได้กลุ่มตัวอย่างจำนวนทั้งสิ้น 300 คน แบ่งเป็นโรงเรียนในกำกับของ สพฐ. 200 คนและโรงเรียนเอกชน 100 คน ทำการแจกแบบสอบถาม 300 ชุดโดยส่งทางไปรษณีย์ ตั้งแต่เดือนมิถุนายนถึงเดือนกันยายน 2557 ได้แบบสอบถามกลับคืนมาทั้งสิ้นจำนวน 195 ชุดหรือคิดเป็นร้อยละ 65 ของแบบสอบถามทั้งหมด

ข้อมูลที่ได้จากแบบสอบถามนำมาวิเคราะห์โดยใช้โปรแกรม SPSS for Windows สถิติที่ใช้ได้แก่ ค่าร้อยละ (Percentage) ค่าเฉลี่ย (Average) และค่าส่วนเบี่ยงเบนมาตรฐาน (S.D.) นำเสนอผลการวิเคราะห์ข้อมูลเป็น 4 ตอน ดังนี้

ตอนที่ 1 ข้อมูลพื้นฐานของผู้ตอบแบบสอบถาม ได้แก่ ตำแหน่งผู้สอน ระดับอาจารย์ ประเภทโรงเรียนที่สังกัด ประสบการณ์การสอน ระดับชั้นและหลักสูตรที่รับผิดชอบ

ตอนที่ 2 การเลือกใช้เครื่องมือและแหล่งสารสนเทศบนอินเทอร์เน็ต เพื่อค้นหาข้อมูลเลิร์นนิงอ็อบเจกต์

ตอนที่ 3 การเลือกใช้ข้อมูลจากแหล่งสารสนเทศบนอินเทอร์เน็ต เพื่อนำมาใช้ในการสร้างเลิร์นนิงอ็อบเจกต์เพื่อนำไปประกอบการเรียนการสอน

ตอนที่ 4 การจัดเก็บเลิร์นนิงอ็อบเจกต์และความต้องการระบบจัดเก็บเลิร์นนิงอ็อบเจกต์

4.1 ผลการวิจัย

ตอนที่ 1 ข้อมูลพื้นฐานของผู้ตอบแบบสอบถาม เป็นข้อมูลรายละเอียดเกี่ยวกับผู้ตอบแบบสอบถามซึ่งเป็นกลุ่มตัวอย่าง ได้แก่ (1) จำนวนผู้ตอบแบบสอบถามจำแนกตามประเภทของโรงเรียน (2) จำนวนผู้ตอบแบบสอบถามจำแนกตามสถานะของผู้สอน (3) จำนวนผู้ตอบแบบสอบถามจำแนกตามประสบการณ์การสอน (4) จำนวนผู้ตอบแบบสอบถามจำแนกตามหลักสูตรที่รับผิดชอบ ดังนี้

ตารางที่ 4.1 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามประเภทของโรงเรียน

ประเภทของโรงเรียนที่สังกัด	จำนวน	ร้อยละ
โรงเรียนในสังกัดของสพฐ.	141	72.3
โรงเรียนเอกชน	54	27.7
รวม	195	100.0

ตารางที่ 4.2 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามสถานะของผู้สอน

สถานะผู้สอน	จำนวน	ร้อยละ
ครูผู้สอน	190	97.4
ครูผู้ช่วยสอน	5	2.6
รวม	195	100

ตารางที่ 4.3 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามประสบการณ์การสอน

ประสบการณ์ในการสอน	จำนวน	ร้อยละ
ไม่ถึง 1 ปี	5	1.5
1-3 ปี	30	15.4
4-6 ปี	34	17.4
7-9 ปี	29	14.9
10 ปีขึ้นไป	97	50.8
รวม	195	100

ตารางที่ 4.4 จำนวนร้อยละของผู้ตอบแบบสอบถามจำแนกตามหลักสูตรที่รับผิดชอบ

หลักสูตรที่รับผิดชอบ	จำนวน	ร้อยละ
หลักสูตรภาษาไทยเท่านั้น	9	4.6
หลักสูตรภาษาอังกฤษเท่านั้น	34	17.4
ทั้งสองหลักสูตร	152	77.9
รวม	195	100

ตารางที่ 4.1 ถึง ตารางที่ 4.4 เป็นข้อมูลพื้นฐานของผู้ตอบแบบสอบถามทั้งสิ้น 195 คน เป็นครูจากโรงเรียนในกำกับดูแลของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) จำนวน 141 คน คิดเป็นร้อยละ 72.3 และครูจากโรงเรียนเอกชนจำนวน 54 คน คิดเป็นร้อยละ 27.7 เป็นครูประจำ/ครูผู้สอน จำนวน 190 คน คิดเป็นร้อยละ 97.4 และเป็นครูผู้ช่วยสอน 5 คน คิดเป็นร้อยละ 2.6

เมื่อจำแนกตามประสบการณ์ของผู้ตอบ พบว่า ร้อยละ 50.8 (97 คน) มีประสบการณ์มากกว่า 10 ปี รองลงมาคือ มีประสบการณ์ 4-6 ปี ร้อยละ 17.4 (34 คน) มีประสบการณ์ 1-3 ปี ร้อยละ 15.4 (30 คน) มีประสบการณ์ 7-9 ปี ร้อยละ 14.9 (29 คน) และน้อยที่สุดคือ มีประสบการณ์ไม่ถึง 1 ปี ร้อยละ 1.5 (5 คน)

เมื่อจำแนกตามหลักสูตรที่รับผิดชอบพบว่า ส่วนใหญ่ร้อยละ 77.9 (152 คน) รับผิดชอบวิชาคณิตศาสตร์ทั้งหลักสูตรภาษาไทยและภาษาอังกฤษ รองลงมาคือ รับผิดชอบเฉพาะหลักสูตรภาษาอังกฤษ ร้อยละ 17.4 (37 คน) และน้อยที่สุดคือ รับผิดชอบเฉพาะหลักสูตรภาษาไทย ร้อยละ 4.6 (9 คน)

ตอนที่ 2 การเลือกใช้เครื่องมือและสารสนเทศบนอินเทอร์เน็ตเพื่อค้นหาข้อมูลเลิร์นนิงอ็อบเจกต์ ได้แก่ (1) แหล่งข้อมูลหลักที่ใช้ในการค้นข้อมูลเพื่อใช้ประกอบการเรียนการสอน (2) เครื่องมือที่นิยมใช้เพื่อชี้แหล่งข้อมูลในการเข้าถึงเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต (3) ข้อมูลที่ใช้ในการสืบค้นเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต (4) ลักษณะของข้อมูลเกี่ยวกับวิชาคณิตศาสตร์ที่ผู้ใช้มักจะค้นหา (5) การนำข้อมูลที่สืบค้นได้ไปใช้เกี่ยวกับการเรียนการสอนวิชา และ(6) การตระหนักรู้ในระดับของปัญหาในการสืบค้นเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต ดังนี้

ตารางที่ 4.5 แหล่งข้อมูลหลักที่ใช้ในการค้นหาข้อมูลเพื่อใช้ประกอบการเรียนการสอน

แหล่งสารสนเทศ	ร้อยละ	อันดับ
หนังสือเรียนตามหลักสูตรแกนกลางของ สพฐ.	27.8	1
สื่อการเรียนรู้หรือเนื้อหาจากอินเทอร์เน็ต	22.8	2
หนังสือเรียนเสริมจากสำนักพิมพ์เอกชนทั่วไป	20.4	3
หนังสือเรียนเสริมจากที่โรงเรียนจัดหาเพิ่ม	17.7	4
สื่อการเรียนการสอนเสริมอื่นๆ เช่น CD ช่วยสอน	11.3	5
รวม	100	

ตารางที่ 4.5 แหล่งข้อมูลหลักที่ใช้ในการค้นหาข้อมูลเพื่อประกอบการเรียนการสอน พบว่าแหล่งข้อมูลที่ผู้ตอบแบบสอบถามใช้มากที่สุดคือ หนังสือเรียนตามหลักสูตรแกนกลางของสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) ร้อยละ 27.8 อันดับ 2 คือ สื่อการเรียนรู้หรือเนื้อหาจากอินเทอร์เน็ต ร้อยละ 22.8 อันดับ 3 คือ หนังสือเรียนเสริมจากภายนอกที่จัดหาเอง ร้อยละ 20.4 อันดับ 4 คือ หนังสือเรียนที่โรงเรียนจัดหาเพิ่มเติม ร้อยละ 17.7 และอันดับสุดท้ายคือ สื่อการเรียนการสอนเสริมอื่นๆ ร้อยละ 11.3

ตารางที่ 4.6 เครื่องมือที่นิยมใช้เพื่อชี้แหล่งข้อมูลในการเข้าถึงเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต

เครื่องใช้ในการชี้แหล่งข้อมูล	ระดับความนิยม	ร้อยละ
เว็บเสิร์ชเอ็นจิน	นิยมมากที่สุด (4.8)	34.0
เว็บไซต์หน่วยงานของรัฐที่เกี่ยวข้องกับการศึกษา	นิยมมาก (3.5)	25.3
เว็บไซต์ของโรงเรียนหรือคลังข้อมูลของโรงเรียน	เป็นที่นิยม (2.9)	20.4

ตารางที่ 4.6 เครื่องมือที่นิยมใช้เพื่อชี้แหล่งข้อมูลในการเข้าถึงเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต

เครื่องใช้ในการชี้แหล่งข้อมูล	ระดับความนิยม	ร้อยละ
เว็บสารบบหรือเว็บไดเร็กทอรี่	เป็นที่นิยม (2.8)	20.3
รวม		100

ตาราง 4.6 เครื่องมือที่นิยมใช้เพื่อชี้แหล่งข้อมูลในการเข้าถึงเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต พบว่า ผู้ตอบแบบสอบถามนิยมใช้เสิร์ชเอ็นจินมากที่สุด ร้อยละ 34 รองลงมาคือ เว็บไซต์หน่วยงานของรัฐที่เป็นเกี่ยวข้องกับการศึกษา เช่น เว็บไซต์ สวทช. เว็บไซต์สสวท. หรือ เว็บไซต์กระทรวงศึกษาธิการ เป็นต้น ร้อยละ 25.3 อันดับสาม คือ เว็บไซต์ของโรงเรียนที่จัดทำขึ้นเองเพื่อให้บริการเลิร์นนิงอ็อบเจกต์แก่ผู้สอนและผู้ที่สนใจ ร้อยละ 20.4 และสุดท้ายคือ เว็บสารบบ (Web Directory) ที่จัดเก็บรายการเลิร์นนิงอ็อบเจกต์ หรือสื่อการสอนแบบเป็นหมวดหมู่ เพื่อเข้าถึงรายการเว็บไซต์ที่จัดเก็บได้โดยตรง ร้อยละ 20.3

จากตารางที่ 4.6 ผู้ตอบแบบสอบถามนิยมใช้เสิร์ชเอ็นจินในการค้นหาเลิร์นนิงอ็อบเจกต์มากที่สุด ทั้งนี้เนื่องจาก เสิร์ชเอ็นจินเป็นเครื่องมือที่ค้นคืนข้อมูลได้หลายประเภท ผู้ค้นไม่จำเป็นต้องรู้แหล่งจัดเก็บเลิร์นนิงอ็อบเจกต์ก็สามารถสืบค้นและเข้าถึงตัวเลิร์นนิงอ็อบเจกต์ผ่านหน้ารายการ ผลการค้นหาได้ทันที สำหรับแหล่งข้อมูลที่ใช้เป็นอันดับสอง คือ เว็บไซต์หน่วยงานของรัฐที่เกี่ยวข้องกับการศึกษา เนื่องจากครูคณิตศาสตร์ระดับมัธยมศึกษามีความมั่นใจในเว็บไซต์ของทางราชการมากกว่าเว็บไซต์ทั่วไป และเนื่องจากเป็นหน่วยงานที่รับผิดชอบทางด้านการศึกษา ยิ่งต้องมีการจัดการเนื้อหาเลิร์นนิงอ็อบเจกต์ที่สอดคล้องกับหลักสูตรแกนกลางฯ มากกว่าเว็บไซต์ทั่วไป ส่วนเว็บไซต์ของโรงเรียนที่จัดทำขึ้นเองก็มักจะไม่เปิดให้ใช้งานระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ได้อย่างเสรี มักใช้ได้เฉพาะบุคลากรภายในหน่วยงานหรือ ภายในโรงเรียนเท่านั้น จึงไม่เป็นที่นิยมสำหรับการค้นคืนข้อมูลของผู้ใช้ทั่วไป ส่วนเว็บสารบบมักจัดระบบเนื้อหาของเว็บทั่วไปตามหมวดหมู่กว้างๆ หรือตามภาพรวมของเว็บไซต์ ไม่สามารถระบุแหล่งจัดเก็บเลิร์นนิงอ็อบเจกต์โดยตรง อีกทั้งการเชื่อมโยงไปยังเว็บต้นทางอาจจะไม่สามารถเข้าใช้ได้กรณีมีการ ทำให้การค้นหาและทดลองเข้าใช้งานต้องใช้เวลามาก

ตารางที่ 4.7 ข้อมูลที่นิยมใช้ในการสืบค้นข้อมูลบนอินเทอร์เน็ต

ข้อมูลที่ใช้ในการสืบค้น	ระดับความนิยม	ร้อยละ
ค้นจากคำสำคัญ	นิยมมาก (4.0)	26.7
ค้นจากชื่อเรื่อง	นิยมมาก (4.0)	26.7

ตารางที่ 4.7 ข้อมูลที่นิยมใช้ในการสืบค้นข้อมูลบนอินเทอร์เน็ต (ต่อ)

ข้อมูลที่ใช้ในการสืบค้น	ระดับความนิยม	ร้อยละ
ค้นจากชื่อหนังสือ หรือชื่อเอกสารการสอน	ไม่เป็นที่นิยม (2.7)	17.8
ค้นจากชื่อบุคคลที่เกี่ยวข้องกับเนื้อหา	ไม่เป็นที่นิยม (2.3)	15.4
ค้นจากเขตข้อมูลหลายๆ เขตข้อมูลพร้อมกัน	ไม่เป็นที่นิยม (2.0)	13.4
รวม		100

จากตารางที่ 4.7 ข้อมูลที่นิยมใช้ในการสืบค้นข้อมูลบนอินเทอร์เน็ต พบว่า ผู้ตอบแบบสอบถามนิยมสืบค้นข้อมูลด้วย คำสำคัญ และชื่อเรื่อง (ชื่อหัวข้อ) มากที่สุดเป็นสองอันดับแรก (ร้อยละ 26.7) ทั้งนี้เนื่องจาก คำสำคัญและชื่อเรื่อง (ชื่อหัวข้อ) สามารถแสดงถึงเนื้อหาของเลิร์นนิ่งอ็อบเจกต์ได้ดีที่สุด สำหรับข้อมูลที่ไม่นิยมนำมาใช้ในการสืบค้นได้แก่ ชื่อหนังสือหรือชื่อเอกสารการสอน (ร้อยละ 17.8) และชื่อบุคคลที่เกี่ยวข้องกับเนื้อหา (ร้อยละ 15.4) เนื่องจากกว้างเกินไป ไม่สามารถระบุถึงเนื้อหาที่ต้องการได้โดยตรง ส่วนการสืบค้นโดยใช้หลายเขตข้อมูลร่วมกัน (การสืบค้นขั้นสูง) ไม่เป็นที่นิยมมากที่สุด (ร้อยละ 13.4) เนื่องจาก การกำหนดขอบเขตที่จำเพาะมากเกินไป หรือควบคุมผลลัพธ์ด้วยเงื่อนไขจำนวนมาก ผู้สอนคาดว่าจะไม่ได้ผลการค้นคืนใด ๆ ออกมาจากระบบเลย และการสืบค้นไม่ได้ต้องการผลการค้นที่เฉพาะเจาะจงมากนัก จึงใช้คำค้นเพียงคำเดียวหรือสองคำ เพื่อให้ได้ผลการค้นในเบื้องต้น จากนั้นจะทำการเลือกรายการที่เกี่ยวข้องอีกครั้ง

ตารางที่ 4.8 ลักษณะของข้อมูลเกี่ยวกับวิชาคณิตศาสตร์ที่มักจะค้นหา

ลักษณะประเภทข้อมูลที่ผู้ใช้ต้องการค้นหา	ระดับการนำไปใช้	ร้อยละ
โจทย์แบบฝึกหัดก่อนเรียน และหลังเรียน	มาก (3.8)	25.4
เนื้อหาบรรยายเกี่ยวกับเรื่องที่สอน	มาก (3.6)	23.8
ข้อสอบหรือแบบทดสอบวัดผลการเรียน	ปานกลาง (3.3)	21.8
กิจกรรมหรือเกมที่เกี่ยวกับเนื้อหารายวิชา	น้อย (2.4)	16.1
รูปภาพประกอบเนื้อหารายวิชา	น้อยมาก (1.9)	12.9
รวม		100

จากตารางที่ 4.8 ลักษณะของข้อมูลเกี่ยวกับวิชาคณิตศาสตร์ที่มักจะค้นหา พบว่า ผู้ตอบแบบสอบถามต้องการสืบค้น โจทย์แบบฝึกหัดก่อนเรียนและหลังเรียนมากที่สุด ร้อยละ 25.4 (ที่ระดับ

มาก 3.8) รองลงมาคือ เนื้อหาบรรยายเกี่ยวกับเรื่องที่สอนมากที่สุด ร้อยละ 23.8 (ที่ระดับมาก 3.6) อันดับสามคือ ข้อสอบหรือแบบทดสอบเพื่อวัดผลการเรียน ร้อยละ 21.8 (ที่ระดับปานกลาง 3.3) ส่วนที่ต้องการน้อย ได้แก่ เกมหรือกิจกรรมที่เกี่ยวข้องกับเนื้อหา ร้อยละ 16.1 (ที่ระดับน้อย 2.4) และรูปภาพประกอบเนื้อหารายวิชามีความต้องการใช้งานในระดับน้อยมาก เพียงร้อยละ 12.9 เท่านั้น

จากตารางที่ 4.8 สรุปได้ว่า ประเภทของข้อมูลที่ผู้สอนต้องการมากคือ ข้อมูลที่สามารถนำไปปรับใช้ได้ทันที ได้แก่ เนื้อหาบรรยายเพื่อนำไปใช้ในการสร้างสื่อการเรียนการสอนได้ทันที หรือนำบางส่วนบางตอนไปปรับใช้ได้ค่อนข้างง่าย ส่วนโจทย์แบบฝึกหัดก่อนเรียนและหลังเรียน และข้อสอบหรือแบบทดสอบวัดผลการเรียน สามารถนำไปเป็นแบบฝึกหัดเสริมให้นักเรียนได้พบโจทย์ปัญหาที่หลากหลายโดยที่ครูผู้สอนไม่ต้องสร้างโจทย์ใหม่ทั้งหมด โดยครูผู้สอนสามารถนำโจทย์ดังกล่าวไปประยุกต์หรือปรับแต่งให้สอนคล้องกับการสอน ทั้งนี้ เพื่อให้ผู้เรียนมีประสบการณ์ในการแก้ปัญหามาก ส่วนเลิร์นนิงอ็อบเจกต์ประเภทเกมหรือ กิจกรรมเสริมนั้น ในระดับมัธยมศึกษาอาจไม่มีความจำเป็น เนื่องจากเนื้อหามีความเป็นวิชาการสูงกว่าระดับประถมศึกษา และมีการนำเสนอเนื้อหาที่เข้าใจได้อยู่แล้ว ส่วนอันดับสุดท้าย ได้แก่ รูปภาพประกอบเนื้อหารายวิชา เนื่องจากมีความเสี่ยงในด้านลิขสิทธิ์เพราะเป็นภาพของผู้อื่น การนำไปใช้อาจจะไม่เหมาะสมนัก

ตารางที่ 4.9 การนำข้อมูลที่สืบค้นคืนได้ไปใช้ในการเรียนการสอนวิชาคณิตศาสตร์

ลักษณะการนำข้อมูลที่สืบค้นได้ไปใช้	ระดับการนำไปใช้	ร้อยละ
ใช้เป็นส่วนเสริมเนื้อหาการสอน เพื่อความหลากหลาย	มาก (4.1)	27.5
ใช้ศึกษาเพิ่มเติมความรู้ในเรื่องใหม่ๆ ของท่านเองเพื่อเทคนิคการสอนที่ดีขึ้น	ปานกลาง (3.5)	23.6
ใช้ประกอบสื่อการเรียนการสอนหลัก	ปานกลาง (3.0)	19.9
ใช้เพื่อเป็นแบบฝึกหัดเสริม หรือการเรียนเสริมนอกเวลาเรียนของนักเรียน	น้อย (2.7)	18.3
ใช้เพื่อเขียนตำราหรือการอ้างอิง	น้อยมาก (1.6)	10.7
รวม		100

ตารางที่ 4.9 การนำข้อมูลที่สืบค้นคืนได้ไปใช้ในการเรียนการสอนวิชาคณิตศาสตร์ พบว่า ผู้ตอบแบบสอบถามนำข้อมูลที่ค้นไปใช้เป็นส่วนเสริมของเนื้อหาการสอนเพื่อความหลากหลายของเนื้อหาวิชามากเป็นอันดับ 1 คิดเป็นร้อยละ 27.5 (ที่ค่าเฉลี่ยระดับมาก 4.1) รองลงมาคือ ใช้ศึกษาเพิ่มเติมความรู้ในเรื่องใหม่ ๆ เพื่อเทคนิคการสอนที่ดีขึ้น คิดเป็นร้อยละ 23.6 (ที่ค่าเฉลี่ยระดับปานกลาง 6.5) อันดับสามคือ

ใช้ประกอบสื่อการเรียนการสอนหลัก คิดเป็นร้อยละ 19.9 (ที่ค่าเฉลี่ยระดับปานกลาง 3.0) ซึ่งใกล้เคียงกับการนำไปใช้เป็นแบบฝึกหัดเสริม หรือการเรียนเสริมนอกเวลาเรียน คิดเป็นร้อยละ 18.3 (ที่ค่าเฉลี่ยระดับน้อย 2.7) และอันดับสุดท้ายคือ นำไปใช้เพื่อเขียนตำราหรือการอ้างอิง ร้อยละ 10.7 (ที่ค่าเฉลี่ยระดับน้อยมาก 1.6) เท่านั้น ทั้งนี้ เนื่องจากการเขียนตำราและการสร้างเอกสารการสอนที่เป็นทางการจะต้องมีการใช้แหล่งข้อมูลที่มีความน่าเชื่อถือ และมีความถูกต้องสูง ทำให้ผู้ใช้ไม่เลือกแหล่งข้อมูลทางอินเทอร์เน็ตเป็นแหล่งอ้างอิงเนื้อหาภายในตัวเล่ม

ตารางที่ 4.10 การตระหนักรู้ระดับของปัญหาในการสืบค้นเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต

ปัญหาที่เกิดจากการสืบค้นข้อมูล	ระดับของปัญหา	ร้อยละ
1. ด้านทักษะในการสืบค้นข้อมูลของผู้ใช้		
1.1 ขาดทักษะภาษาอังกฤษเพื่อสืบค้น	มาก (4.5)	12.0
1.2 ขาดความรู้เกี่ยวกับแหล่งที่ใช้สืบค้น	ปานกลาง (3.6)	4.0
1.3 ขาดความรู้การเลือกใช้คำค้นที่เหมาะสม	ปานกลาง (3.3)	3.9
1.4 ขาดความชำนาญในการสืบค้น	น้อย (2.0)	2.8
2. ด้านปัญหาในการใช้งานระบบจัดเก็บเลิร์นนิงอ็อบเจกต์		
2.1 ข้อมูลมีลิขสิทธิ์หรือมีค่าใช้จ่าย	มากที่สุด (4.7)	9.5
2.2 ผลลัพธ์มีความซ้ำซ้อนหรือเหมือนกันในหลายรายการของผลการสืบค้น	มาก (4.1)	7.2
2.3 ชื่อเรื่องไม่สอดคล้องกับเนื้อหา หรือเป็นการหลอกล่อเพื่อเข้าไปยังเว็บที่เป็นเนื้อหาอื่นๆ	ปานกลาง (3.0)	7.0
2.4 ไม่ระบุแหล่งที่มาหรือดูไม่น่าเชื่อถือ	ปานกลาง (3.0)	7.0
2.5 รายชื่อสื่อ (ชื่อเรื่อง) ที่เป็นผลลัพธ์การสืบค้นไม่ปรากฏคำค้นที่ใช้ ทำให้มีผลกับการตัดสินใจเลือกใช้ผลการค้นคืนแต่ละรายการ	น้อย (2.9)	6.8
2.6 คำอธิบายของผลลัพธ์ (บทคัดย่อ) แต่ละรายการไม่ปรากฏคำค้นที่ใช้ ทำให้มีผลกับการตัดสินใจเลือกใช้ผลการค้นคืนแต่ละรายการ	น้อย (2.9)	6.8
2.7 รายการผลลัพธ์มีจำนวนหน้าที่มากเกินไป	น้อยมาก (1.9)	6.7
2.8 ไม่มีการแนะนำคำค้นที่เกี่ยวข้อง	น้อยมาก (1.8)	6.6

ตารางที่ 4.10 การตระหนักรู้ระดับของปัญหาในการสืบค้นเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต

ปัญหาที่เกิดจากการสืบค้นข้อมูล	ระดับของปัญหา	ร้อยละ
2. ด้านปัญหาในการใช้งานระบบจัดเก็บเลิร์นนิงอ็อบเจกต์		
2.9 ข้อมูลมีเนื้อหาไม่ถูกต้อง	น้อย (2.8)	6.5
2.10 ไม่มีการเชื่อมโยงไปยังเว็บไซต์หรือแหล่งอื่นๆ เพิ่ม	น้อย (2.8)	6.5
2.11 ไม่มีคำวิจารณ์หรือการให้คะแนนรายการผลการค้นคืนแต่ละรายการ	น้อย (2.5)	3.8
รวม		100

ตารางที่ 4.10 การตระหนักรู้ระดับของปัญหาที่เกิดขึ้นในการสืบค้นเลิร์นนิงอ็อบเจกต์ แบ่งเป็น 2 ด้าน ดังนี้

1) ปัญหาด้านทักษะการสืบค้นข้อมูลของผู้ใช้ พบว่า สิ่งที่เป็นปัญหาสำหรับการสืบค้นมากที่สุดคือ การขาดทักษะภาษาอังกฤษในการสืบค้นเนื้อหารายวิชาเป็นภาษาอังกฤษ คิดเป็นร้อยละ 12 (ที่ค่าเฉลี่ยระดับมาก 4.5) เนื่องจากคำศัพท์ที่ใช้เป็นคำศัพท์เฉพาะทางคณิตศาสตร์ วิธีการสะกดคำและประโยคบริบทอาจทำให้เกิดความเข้าใจผิดพลาด หรือถ้าต้องการสืบค้นเลิร์นนิงอ็อบเจกต์จากแหล่งข้อมูลภาษาต่างประเทศ การเลือกคำค้นที่เป็นภาษาอังกฤษที่ไม่ครบถ้วนถูกต้องอาจทำให้ ผลการค้นเกิดความคลาดเคลื่อนหรือไม่ได้ผลการค้น อันดับรองลงมาคือ การขาดความรู้ในแหล่งที่ใช้ค้นคืนและเข้าถึงเลิร์นนิงอ็อบเจกต์ คิดเป็นร้อยละ 4.0 (ที่ค่าเฉลี่ยระดับ ปานกลาง 3.6) ซึ่งใกล้เคียงกับประเด็นปัญหาการเลือกใช้คำค้นที่เหมาะสม คิดเป็นร้อยละ 3.9 (ที่ค่าเฉลี่ยระดับปานกลาง 3.3) ส่วนประเด็นที่มีปัญหาระดับน้อยคือ การขาดความชำนาญในการค้น เช่น การเลือกใช้เครื่องมือ การเลือกคำสั่งการค้นคืน และการแสดงผล คิดเป็นร้อยละ 2.8 (ที่ค่าเฉลี่ยระดับน้อย 2.0)

2) ด้านปัญหาในการใช้งานระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ พบว่า ประเด็นที่ผู้ตอบแบบสอบถามเห็นว่าเป็นปัญหามากที่สุดคือ ข้อมูลที่ต้องการมีค่าใช้จ่ายหรือมีลิขสิทธิ์ ร้อยละ 9.5 (ที่ค่าเฉลี่ยระดับมากที่สุด 4.7) รองลงมาคือ รายการผลการค้นมีความซ้ำซ้อนหรือมีความเหมือนกัน คิดเป็นร้อยละ 7.2 (ที่ค่าเฉลี่ยระดับมาก 4.1) อันดับที่สามคือ ปัญหาชื่อเรื่องของเว็บไซต์ไม่สอดคล้องกับเนื้อหา หรือเป็นการหลอกล่อเพื่อเข้าไปยังเว็บที่เป็นเนื้อหาอื่นๆ และปัญหาการไม่ระบุแหล่งที่มา หรือดูไม่น่าเชื่อถือ คิดเป็นร้อยละ 7 (ที่ค่าเฉลี่ยระดับปานกลาง 3.0) ส่วนประเด็นที่ผู้ตอบแบบสอบถามเห็นว่ามีปัญหาในระดับน้อย ได้แก่ รายชื่อสื่อ (ชื่อเรื่อง) ที่เป็นผลลัพธ์การสืบค้นไม่ปรากฏคำค้นที่ใช้ และคำอธิบายของผลลัพธ์ (บทคัดย่อ) แต่ละรายการไม่ปรากฏคำค้นที่

ใช้ ทำให้มีผลกับการตัดสินใจเลือกใช้ผลการค้นคืนแต่ละรายการ คิดเป็นร้อยละ 6.8 (ที่ค่าเฉลี่ยระดับน้อย 2.9) ถัดมาคือ ข้อมูลมีเนื้อหาไม่ถูกต้อง และไม่มีการเชื่อมโยงไปยังเว็บไซต์หรือแหล่งอื่นๆ เพิ่มคิดเป็นร้อยละ 6.5 (ที่ค่าเฉลี่ยระดับน้อย 2.8) ส่วนประเด็นปัญหาที่ผู้ตอบแบบสอบถามเห็นว่ามีปัญหาในระดับน้อยมาก (ที่ค่าเฉลี่ย 1.9 และ1.8) คิดเป็นร้อยละ 6.7 และ 6.6 ได้แก่ ผลการค้นคืนที่มีจำนวนหลายหน้าหรือมีจำนวนมากเกินไป และไม่มีการแนะนำคำค้นที่เกี่ยวข้อง

ตอนที่ 3 การเลือกใช้ข้อมูลจากแหล่งสารสนเทศบนอินเทอร์เน็ตเพื่อนำตัวสื่อการสอนและเลิร์นนิงอ็อบเจกต์ไปประกอบการเรียนการสอน ได้แก่ ข้อมูลที่ใช้ในการพิจารณาเพื่อเปิดใช้รายการผลการสืบค้นจากหน้ารวมผลการค้น และการคัดกรอง หรือปรับแต่งการแสดงผลการสืบค้น

ตารางที่ 4.11 ข้อมูลที่ใช้ในการพิจารณาเพื่อเปิดใช้รายการผลการสืบค้นจากหน้ารวมผลการค้น

เขตข้อมูล	ระดับการใช้	ร้อยละ
1. คำสำคัญที่ปรากฏบนชื่อเรื่อง (Keyword)	มากที่สุด (4.7)	7.0
2. ประเภทของไฟล์ (Word, PDF, Excel เป็นต้น)	มากที่สุด (4.5)	6.9
3. ชื่อเรื่องเต็ม (Title)	มากที่สุด (4.5)	6.8
4. รายละเอียดเนื้อหา ส่วนประกอบภายในชิ้นงาน (Abstract/ Description)	มาก (4.3)	6.8
5. หัวเรื่อง/ประเด็นหลักของเนื้อหา (Subject Heading)	มาก (3.8)	6.8
6. ภาษาที่ใช้ในชิ้นงาน (Language)	มาก (3.6)	6.3
7. แหล่งที่มาของข้อมูลในเนื้อหา (Source)	มาก (3.6)	6.3
8. ระดับความยากง่ายของเนื้อหา	มาก (3.9)	6.3
9. ชื่อหน่วยงานหรือเว็บไซต์ที่เผยแพร่ข้อมูล (Publisher)	มาก (3.5)	6.1
10. ข้อมูลแสดงสิทธิ์ในการเข้าใช้ หรือแสดงลิขสิทธิ์ (Copyright)	ปานกลาง (3.3)	6.0
11. วันที่/ปีที่ผลิตและเผยแพร่เอกสาร (Date)	ปานกลาง (3.4)	6.0
12. แหล่งข้อมูลเชื่อมโยงอื่นๆ (Related)	ปานกลาง (3.4)	6.0
13. ชื่อผู้แต่งหรือผู้สร้างผลงาน (Author)	ปานกลาง (3.4)	5.9
14. ลำดับการแสดงผลรายการก่อน-หลัง	ปานกลาง (3.4)	5.7
15. คำวิจารณ์หรือการให้คะแนน (Review)	ปานกลาง (3.3)	5.7
16. รหัสระบุตามมาตรฐานสากล เช่น URI ISBN เป็นต้น	น้อย (2.2)	5.3
รวม		100

จากตารางที่ 4.11 ข้อมูลที่ใช้ในการพิจารณาเพื่อเปิดใช้รายการผลการสืบค้นจากหน้ารวมผลการค้นพบว่า ร้อยละของการเลือกข้อมูลแต่ละข้อมีความใกล้เคียงกันคือระหว่า 5.3- 7.0 โดยข้อมูลที่มีค่าเฉลี่ยการใช้ในระดับมากที่สุดมี 3 ราย ที่ค่าคะแนน 4.7 4.5 และ 4.5 ได้แก่ มีคำสำคัญปรากฏบนชื่อเรื่อง ชื่อเรื่อง และประเภทของไฟล์

ส่วนข้อมูลที่ใช้ในระดับมาก ประกอบด้วย 6 รายการ ได้แก่ รายละเอียดเนื้อหาส่วนประกอบภายในชิ้นงาน (ค่าเฉลี่ย 4.3) ระดับความยากง่ายของเนื้อหา (ค่าเฉลี่ย 3.9) หัวเรื่อง/ประเด็นหลักของเนื้อหา (ค่าเฉลี่ย 3.8) ภาษาที่ใช้ในชิ้นงาน และ แหล่งที่มาของข้อมูลในเนื้อหา (ค่าเฉลี่ย 3.6) ชื่อหน่วยงานหรือเว็บไซต์ที่เผยแพร่ข้อมูล (ค่าเฉลี่ย 3.5)

ข้อมูลที่ใช้ประกอบพิจารณาในระดับปานกลาง ประกอบด้วย 6 รายการได้แก่ วันที่หรือปีที่ผลิตและเผยแพร่เอกสาร แหล่งข้อมูลเชื่อมโยงอื่นๆ ชื่อผู้แต่งหรือผู้สร้างผลงาน ลำดับการแสดงผลรายการก่อน-หลัง (ค่าเฉลี่ย 3.4) ข้อมูลแสดงสิทธิ์ในการเข้าใช้ หรือแสดงลิขสิทธิ์ และคำวิจารณ์หรือการให้คะแนน (ค่าเฉลี่ย 3.3)

ส่วนข้อมูลที่ใช้ประกอบการพิจารณาคัดเลือกในระดับน้อยมีเพียง 1 รายการ คือ เลขรหัสระบุเลิร์นนิงอ็อบเจกต์ตามมาตรฐานสากล (ค่าเฉลี่ย 2.2)

ตารางที่ 4.12 การคัดกรอง หรือปรับแต่งการแสดงผลการสืบค้น

การคัดกรองข้อมูล	ระดับการใช้	ร้อยละ
1. แสดงเฉพาะรายการที่เป็นฉบับเต็ม	มากที่สุด (4.5)	22.6
2. การจัดกลุ่มเนื้อหาตามระดับชั้นปีการศึกษา	มาก (3.9)	19.9
3. การจัดกลุ่มเนื้อหาตามประเด็นเรื่อง	มากที่สุด (4.6)	22.2
4. แสดงเฉพาะประเภทของชิ้นงานที่ต้องการ เช่น เป็นแบบฝึกหัด เอกสารบรรยายเนื้อหา ข้อสอบ การบ้าน เป็นต้น	มาก (3.9)	19.8
5. การแสดงรูปภาพหรือสื่ออื่นๆ ที่อยู่ในเนื้อหาไว้ในรายการสืบค้น	ปานกลาง (3.1)	15.6
รวม		100

จากตารางที่ 4.12 การคัดกรอง หรือปรับแต่งการแสดงผลการสืบค้น พบว่า ผู้ตอบแบบสอบถามใช้การแสดงเฉพาะรายการที่เป็นฉบับเต็มมากที่สุด รองลงมาคือ การจัดกลุ่มเนื้อหาตามประเด็นเรื่อง ที่ร้อยละ 22.6 และ 22.2 (ระดับการใช้มากที่สุด 4.5 และ 4.6) ตามลำดับ รองลงมาคือ การจัดกลุ่มเนื้อหาแบบตามระดับชั้นปีการศึกษา และการแสดงเฉพาะประเภทชิ้นงานที่ต้องการ

ที่ร้อยละ 19.9 และ 19.8 (ระดับการใช้มาก 3.9) ตามลำดับ และเงื่อนไขที่ใช้น้อยที่สุดคือ การแสดงรูปภาพ หรือสื่ออื่นๆ ที่อยู่ในเนื้อหาไว้ในรายการผลการสืบค้น ร้อยละ 15.6 (ระดับการใช้ปานกลาง 3.1)

ตอนที่ 4 การจัดเก็บเลิร์นนิงอ็อบเจกต์ เพื่อวิเคราะห์ถึงความสามารถในการจัดการเลิร์นนิงอ็อบเจกต์ และศึกษาความต้องการของในการจัดการเลิร์นนิงอ็อบเจกต์ที่สร้างขึ้นเอง

ตารางที่ 4.13 วิธีการจัดเก็บเลิร์นนิงอ็อบเจกต์ที่มีอยู่เดิมหรือสร้างขึ้นใหม่

วิธีการจัดเก็บเลิร์นนิงอ็อบเจกต์	ระดับการใช้	ร้อยละ
1. จัดเก็บในคอมพิวเตอร์ส่วนตัวไม่มีการเผยแพร่	มากที่สุด (4.8)	33.7
2. ไม่มีการจัดเก็บไฟล์ (ใช้แล้วทิ้ง) แต่พิมพ์ออกมาเป็นเอกสารเพื่อใช้งาน	ปานกลาง (3.4)	23.6
3. ใช้โปรแกรมที่สถานบันกำหนด CMS หรือ คลังข้อมูลส่วนกลาง	มาก (3.6)	25.4
4. ใช้โปรแกรมสำเร็จรูปในการจัดเก็บเป็นระบบส่วนตัว	น้อย (2.4)	17.3
รวม		100

จากตารางที่ 4.13 วิธีการจัดเก็บเลิร์นนิงอ็อบเจกต์ที่มีอยู่เดิมหรือสร้างขึ้นใหม่ พบว่า วิธีการที่ผู้ตอบแบบสอบใช้มากที่สุดในการจัดเก็บเลิร์นนิงอ็อบเจกต์คือ จัดเก็บไว้ในคอมพิวเตอร์ส่วนตัวไม่มีการเผยแพร่ ที่ร้อยละ 33.7 (ระดับการใช้มากที่สุด ค่าเฉลี่ย 4.8) รองลงมาคือ ใช้โปรแกรมที่สถาบันกำหนด คือระบบจัดการการเรียนรู้ (CMS) ที่ร้อยละ 25.4 (ระดับการใช้มาก ค่าเฉลี่ย 3.6) อันดับสามคือ ไม่มีการจัดเก็บไฟล์หรือใช้แล้วทิ้งเป็นครั้งๆ อยู่ที่ร้อยละ 23.6 (ระดับการใช้ปานกลาง ค่าเฉลี่ย 3.4) ส่วนวิธีการที่ใช้น้อยที่สุดคือ ใช้โปรแกรมสำเร็จรูปในการจัดเก็บเป็นระบบส่วนตัว ร้อยละ 17.3 (ระดับการใช้น้อย ค่าเฉลี่ย 2.4)

สำหรับความต้องการระบบจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ จะใช้ข้อคำถามปลายเปิดรวบรวมความต้องการและข้อเสนอแนะได้ ดังนี้

1.) อยากให้มีแหล่งข้อมูลที่เปิดเผยและใช้งานได้ฟรี สามารถเข้าถึงได้สะดวก

2.) อยากให้มีแหล่งรวบรวมสื่อที่เป็นของคนไทยและสอดคล้องกับการเรียนการสอนมากกว่านี้

3.) ควรมีการรวมกลุ่ม หรือหน่วยงานกลางที่รับผิดชอบการจัดสร้าง จัดเก็บและเผยแพร่อย่างเป็นระบบทุกโรงเรียนสามารถมีส่วนร่วมและเข้าถึงได้

4.) ปัจจุบันแหล่งข้อมูลมีการปิดกัน และเสียค่าใช้จ่ายมาก และการจัดซื้อเองอาจได้เนื้อหาที่ไม่ตรงกับหลักสูตร จึงอยากให้มีแหล่งรวบรวมสื่อการเรียนรู้ที่เข้าถึงได้แบบไม่เสียค่าใช้จ่าย

4.2 สรุปและอภิปรายผล

จากการศึกษาพฤติกรรมการสืบค้นเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตของครูคณิตศาสตร์จากโรงเรียนในสังกัดของสพฐ. และโรงเรียนเอกชนที่มีการเรียนการสอนในหลักสูตรสองภาษา โดยกำหนดวัตถุประสงค์ของการศึกษาเพื่อ (1) ศึกษาแหล่งข้อมูลที่ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษา หลักสูตรสองภาษาใช้ในการค้นคืนเลิร์นนิงอ็อบเจกต์ (2) วิธีการค้นคืนและคัดเลือกรายการเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต (3) การนำเลิร์นนิงอ็อบเจกต์ที่ได้ไปใช้ในการเรียนการสอน (4) ปัญหาที่พบในการค้นคืนเลิร์นนิงอ็อบเจกต์ ได้แก่ ปัญหาจากตัวบุคคล และปัญหาจากเครื่องมือและแหล่งจัดเก็บข้อมูล รวมทั้งปัญหาของข้อมูลที่ใช้ในการสืบค้นเลิร์นนิง-อ็อบเจกต์ และ (5) วิธีการจัดเก็บเลิร์นนิงอ็อบเจกต์ โดยผลการศึกษาในส่วนนี้จะนำไปใช้เป็นข้อมูลพื้นฐานในการการออกแบบโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ ทั้งนี้เพื่อให้รายละเอียดในการจัดเก็บและการแสดงผลการค้นสามารถตอบสนองกับความต้องการของผู้ใช้มากที่สุด โดยสรุปผลการศึกษาได้ ดังนี้

4.2.1 แหล่งข้อมูลหลักที่ใช้ในการค้นหาข้อมูลเพื่อใช้ประกอบการเรียนการสอน

ผลการศึกษา พบว่า ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษาใช้หนังสือเรียนตามหลักสูตรแกนกลางของสพฐ. เป็นแหล่งข้อมูลหลักในการเตรียมเอกสารการเรียนการสอนมากที่สุดและใช้เสิร์ชเอ็นจินสืบค้นข้อมูลเพิ่มเติมจากอินเทอร์เน็ต เนื่องจากสามารถเข้าถึงข้อมูลจากแหล่งอื่นได้รวดเร็ว รวมถึงได้ข้อมูลในรูปแบบดิจิทัลที่สามารถนำมาใช้งานได้สะดวกกว่าการสร้างเลิร์นนิงอ็อบเจกต์ขึ้นใหม่ ส่วนแหล่งข้อมูลอื่น เช่น หนังสือเรียนเสริม หรือหนังสือเรียนที่โรงเรียนจัดหามาให้ ตลอดจนสื่อการเรียนการสอนแบบสื่อประสมนั้นมีข้อจำกัดคือ ด้านงบประมาณ บางโรงเรียนไม่มีการจัดหามาได้ นอกจากนี้ยังเห็นว่า การนำข้อมูลจากหนังสือหรือสื่อที่มีลิขสิทธิ์มาใช้งานนั้นอาจจะไม่สะดวกเท่าการนำเลิร์นนิงอ็อบเจกต์ที่มีผู้สร้างไว้เรียบร้อยแล้วและเผยแพร่ในระบบเปิดแบบเสรีมาใช้งาน

อย่างไรก็ดี ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษา ยังคมนิยมใช้ข้อมูลจากแหล่งข้อมูลที่น่าเชื่อถือ คือ หน่วยงานทางการศึกษาของภาครัฐ อาทิ กระทรวงศึกษาธิการ

สำนักงานส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี เป็นต้น เนื่องจากเป็นหน่วยงานที่รับผิดชอบหลักสูตรโดยตรง ทำให้สามารถหาเลิร์นนิงอ็อบเจกต์ได้ตรงตามเนื้อหาของหลักสูตรการเรียนการสอนมากที่สุด แต่จากการศึกษาพบว่า เว็บไซต์ของสถาบันทางการศึกษายังมีการจัดทำและสร้างเลิร์นนิงอ็อบเจกต์เพื่อเผยแพร่ให้งานในจำนวนน้อยมากและมีเพียงบางวิชา บางหัวเรื่องเท่านั้น

4.2.2 วิธีการสืบค้นและคัดเลือกรายการเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ตเพื่อนำไปใช้ในการเรียนการสอน

จากการศึกษาพบว่า วิธีการสืบค้นที่ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษาใช้มากที่สุดคือ การใส่คำสำคัญที่เกี่ยวข้องกับเรื่องที่ต้องการค้น โดยมักใช้คำค้นที่เป็นคำสำคัญในเนื้อหา พอๆ กับชื่อเรื่องของบทเรียน เพราะเชื่อว่าจะได้ผลการค้นคืนที่เกี่ยวข้องกับเรื่องที่ต้องการมากที่สุด ส่วนข้อมูลที่ไม่นิยมใช้ในการสืบค้นคือ ชื่อหนังสือ และชื่อบุคคลที่เกี่ยวข้องกับเนื้อหานั้น เนื่องจากจะได้ข้อมูลที่กว้างเกินไป ส่วนการสืบค้นจากข้อมูลหลายตัว หรือการค้นขั้นสูง (Advance Search) ไม่เป็นที่นิยม เนื่องจากผลการค้นคืนจะเฉพาะเจาะจงมากเกินไป และอาจจะไม่มีรายการที่ตรงกับเงื่อนไขออกมาเลย และการกำหนดเงื่อนไขมากเกินไปอาจทำให้ผลการค้นจากเรื่องอื่นๆ ที่อาจเกี่ยวข้องถูกตัดออกไปด้วย

สำหรับเนื้อหาที่ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษานิยมสืบค้นได้แก่ บทบรรยายและเนื้อหาการสอนหลัก เนื่องจากไม่จำเป็นต้องถอดเนื้อหาจากหนังสือหรือสิ่งพิมพ์ สามารถนำข้อความเนื้อหาที่เป็นแบบดิจิทัลไปใช้งานได้ทันที นอกจากนี้ ยังนิยมค้นหาแบบฝึกหัด ข้อสอบ และโจทย์ปัญหาเพื่อนำไปใช้เป็นแบบฝึกหัดที่มีความหลากหลาย และสามารถดัดแปลงให้มีความซับซ้อน แปลกใหม่ได้ง่าย โดยไม่ต้องเริ่มออกข้อสอบเองตั้งแต่ต้น ส่วนภาพประกอบ เกม และกิจกรรม ไม่เป็นที่นิยมในการสืบค้น เนื่องจาก ประเด็นปัญหาด้านลิขสิทธิ์และการนำมาดัดแปลงเพื่อใช้งานนั้นทำได้ยาก

4.2.3 การนำเลิร์นนิงอ็อบเจกต์ที่ได้เพื่อไปใช้ในการเรียนการสอน

จากการศึกษาพบว่า เลิร์นนิงอ็อบเจกต์ที่สืบค้นคืนได้นั้น ส่วนใหญ่ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษาจะนำไปใช้เป็นส่วนประกอบของเนื้อหาการสอนแบบเสริม กล่าวคือ เนื้อหาที่เพิ่มเติมจากเนื้อหาหลัก หรือเอกสารประกอบการเรียนการสอนหลัก เช่น แบบฝึกหัดเพิ่มเติม แบบทดสอบนอกบทเรียน เป็นต้น โดยส่วนเสริมเหล่านี้ครูผู้สอนต้องเป็นผู้สร้างขึ้นเอง เพิ่มเติมจากหนังสือเรียนตามหลักสูตรแกนกลาง นอกจากนี้ ยังมีบางส่วนที่นำไปใช้

ในการสอนหลัก กรณีที่เนื้อหาหลักมีรายละเอียดไม่เพียงพอ หรือบางครั้งผู้ใช้ต้องการตัวอย่างที่แตกต่างจากตัวอย่างในหนังสือเรียนตามปกติ แต่ส่วนการเขียนตำราหรือเอกสารการสอนนั้นผู้ใช้ไม่นิยมนำข้อมูลจากอินเทอร์เน็ตมาใช้ประกอบในเนื้อหาตำรามากนัก เนื่องจากต้องมีรายละเอียดที่เป็นทางการ แต่จะใช้หนังสือเรียนตามหลักสูตรเป็นเอกสารหลักในการเลือกเนื้อหาและบทบรรยายมากกว่า

4.2.4 ปัญหาที่พบในการสืบค้นเลิร์นนิงอ็อบเจกต์

การตระหนักรู้ปัญหาด้านตระหนักรู้ปัญหาในการสืบค้นข้อมูล

(1) ด้านทักษะการสืบค้นข้อมูลของผู้ใช้ จากการศึกษาพบว่า ปัญหาที่ครูคณิตศาสตร์ระดับมัธยมศึกษา หลักสูตรสองภาษาพบมากที่สุดคือ ปัญหาการใช้ภาษาอังกฤษในการสืบค้นข้อมูล แม้ศัพท์คณิตศาสตร์จะเป็นคำศัพท์เฉพาะและศัพท์คณิตศาสตร์ระดับมัธยมศึกษายังมีไม่มากและไม่ซับซ้อน ซึ่งสอดคล้องกับผลการศึกษาที่พบว่า การใช้คำค้นที่เหมาะสมเป็นปัญหาหนึ่งในการสืบค้นเลิร์นนิงอ็อบเจกต์ทางคณิตศาสตร์ โดยเฉพาะกรณีที่ชื่อของเอกสารไม่มีคำสำคัญปรากฏอยู่

(2) ด้านการใช้งานระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ พบว่า ประเด็นที่ผู้ตอบแบบสอบถามเห็นว่าเป็นปัญหามากที่สุดคือ ข้อมูลที่ต้องการมีค่าใช้จ่ายหรือมีลิขสิทธิ์ ซึ่งทำให้ไม่สามารถนำไปใช้ได้อย่างเสรี หรือถูกปิดกั้นการเข้าถึง รองลงมาคือ รายการผลการค้นมีความซ้ำซ้อนหรือมีความเหมือนกัน หรือ ผลการสืบค้นมีหลายรายการแต่เชื่อมโยงไปยังจุดหมายเดียวกัน ซึ่งเป็นลักษณะของผลการสืบค้นจากเสิร์ชเอ็นจิน ทำให้ผู้ค้นเกิดความสับสนและใช้เวลาค้นหาหน้าเอกสารใหม่ๆ นานขึ้น ส่วนข้อที่ผู้ตอบแบบสอบถามเห็นว่าเป็นปัญหาน้อยมาก ได้แก่ รายการผลลัพธ์มีจำนวนหน้าที่มากเกินไป และไม่มีการแนะนำคำค้นที่เกี่ยวข้อง ซึ่งขัดแย้งกับประเด็นก่อนหน้าก่อนหน้าที่ผู้ตอบแบบสอบถามเห็นว่า ผลการค้นที่ที่มีจำนวนมากเกินไปทำให้ต้องใช้เวลานานในการพิจารณาคัดเลือก และการกำหนดคำค้นที่เกี่ยวข้องเป็นปัญหาในการสืบค้น

ข้อมูลที่ครูคณิตศาสตร์ระดับมัธยมศึกษาใช้เพื่อคัดเลือกรายการผลการค้นเพื่อนำเลิร์นนิงอ็อบเจกต์ไปใช้

ผลการศึกษาพบว่า ข้อมูลที่ผู้ตอบแบบสอบถามนำมาใช้เพื่อพิจารณาคัดเลือกผลการสืบค้นที่ต้องการ แบ่งออกเป็น 3 กลุ่มคือ (1) ข้อมูลที่ใช้ในการตัดสินใจอันดับแรก (หรือการตัดสินใจในเบื้องต้น) เน้นที่เนื้อหาและรูปแบบของเอกสาร (2) ข้อมูลที่ใช้ประกอบการตัดสินใจเพื่อจำแนกรายละเอียดด้านเนื้อหา (3) ข้อมูลเกี่ยวกับสิทธิในการใช้งานเลิร์นนิงอ็อบเจกต์

(1) ข้อมูลที่ใช้ในการตัดสินใจอันดับแรก (หรือการตัดสินใจในเบื้องต้น) เน้นที่เนื้อหาและรูปแบบของเอกสาร ได้แก่ ข้อมูลที่มีการใช้ในระดับมากที่สุด ประกอบด้วย ชื่อเรื่อง ประเภทของไฟล์ ทั้งนี้เนื่องจากชื่อเรื่องสามารถบ่งบอกถึงเนื้อหา และประเภทของไฟล์บอกถึงความสามารถในการนำไปใช้ เป็นข้อมูลที่ใช้ในการตัดสินใจในเบื้องต้น

(2) ข้อมูลที่ใช้ประกอบการตัดสินใจเพื่อจำแนกรายละเอียดด้านเนื้อหา ได้แก่ กลุ่มข้อมูลที่มีการใช้ในระดับมาก ประกอบด้วย รายละเอียดเนื้อหา หัวเรื่อง ภาษา แหล่งที่มาของข้อมูล ระดับความยากง่ายของเนื้อหา และหน่วยงานที่เผยแพร่ ข้อมูลในกลุ่มนี้จะแสดงถึง รายละเอียดที่มากขึ้นของเนื้อหา ทั้งในด้านความเกี่ยวข้อง ความยากง่าย และความน่าเชื่อถือของเนื้อหา เป็นกลุ่มข้อมูลเพิ่มเติมที่ใช้ประกอบการตัดสินใจเลือกใช้เลิร์นนิงอ็อบเจกต์ที่เกี่ยวข้อง

(3) ข้อมูลเกี่ยวกับสิทธิในการใช้งานเลิร์นนิงอ็อบเจกต์ ได้แก่ กลุ่มข้อมูลที่มีการใช้ในระดับปานกลาง ประกอบด้วย สิทธิการเข้าใช้ วันที่เผยแพร่ ชื่อผู้แต่ง เอกสารที่เกี่ยวข้อง และคำวิจารณ์หรือการให้คะแนน ข้อมูลในกลุ่มนี้จะเน้นในเรื่องของสิทธิในการนำข้อมูลไปใช้เป็นหลัก ทั้งนี้เนื่องจากเลิร์นนิงอ็อบเจกต์สามารถนำไปใช้ได้ทั้งในลักษณะของการนำไปใช้ซ้ำ การนำไปปรับปรุงแก้ไข หรือการนำองค์ประกอบบางส่วนไปใช้

ส่วนข้อมูลที่มีระดับการใช้น้อยหรือแทบไม่ใช้ คือ รหัสประจำตัวเอกสาร (เช่น URI ISBN เป็นต้น) ซึ่งเป็นข้อมูลในการระบุตัวตน หรือแหล่งที่อยู่ของเอกสารบนเว็บ ไม่ได้มีความเกี่ยวข้องกับเนื้อหา จึงมีการนำมาใช้ในการพิจารณาคัดเลือกเอกสารในระดับน้อย

สำหรับการจัดกลุ่มผลการค้น ผู้ตอบแบบสอบถามต้องการให้มีการจัดกลุ่มผลการค้นตามประเภทเนื้อหา และรูปแบบข้อมูลมากที่สุด รองลงมาคือ ระดับชั้นของการศึกษาและประเภทของสื่อการเรียนรู้ (เช่น เป็นแบบฝึกหัด เอกสารบรรยายเนื้อหา ข้อสอบ การบ้าน เป็นต้น)

วิธีการจัดเก็บเลิร์นนิงอ็อบเจกต์

จากการศึกษาพบว่า ผู้ตอบแบบสอบถามส่วนใหญ่ถึงร้อยละ 76.4 มีการจัดเก็บเลิร์นนิ่งอ็อบเจกต์ทั้งที่สืบค้นได้และทั้งที่สร้างขึ้นมาใหม่ โดยกลุ่มหนึ่งเก็บไว้ในคอมพิวเตอร์ส่วนตัว (ไม่เผยแพร่) กลุ่มหนึ่งจัดเก็บไว้ในระบบจัดการเลิร์นนิ่งอ็อบเจกต์ของโรงเรียน (หากมี) กลุ่มหนึ่งจัดเก็บไว้ในโปรแกรมสำเร็จรูปในการจัดเก็บเป็นระบบส่วนตัว ทั้งนี้เพื่อเก็บไว้ใช้งานในอนาคต ขณะที่ผู้ชายแบบสอบถาม ร้อยละ 23.6 ไม่ได้จัดเก็บในรูปของแฟ้มข้อมูลอิเล็กทรอนิกส์ แต่พิมพ์ออกมาเก็บไว้ในรูปของสิ่งพิมพ์ ซึ่งพฤติกรรมดังกล่าวแสดงให้เห็นว่าครูผู้สอนมีความต้องการในการจัดเก็บสื่อการเรียนรู้หรือเลิร์นนิงอ็อบเจกต์ไว้สำหรับนำกลับมาใช้อีก ซึ่งสอดคล้องกับข้อเสนอแนะด้านความต้องการระบบการจัดเก็บที่ผู้ตอบแบบสอบถามต้องการให้มีหน่วยงานกลาง

หรือการรวมกลุ่มเพื่อนำเลิร์นนิงอ็อบเจกต์ของไทยมาจัดเก็บไว้ด้วยกัน ทั้งนี้เพื่อความสะดวกในสืบค้นข้อมูล และสามารถแบ่งปันให้ใช้งานร่วมกันได้โดยไม่เสียค่าใช้จ่าย อย่างไรก็ดี ในการนำเลิร์นนิงอ็อบเจกต์ของบุคคลอื่นไปใช้ควรมีการขออนุญาตจากเจ้าของ หรือมีการอ้างอิงเพื่อให้เกียรติแก่เจ้าของผลงาน

4.2.5 การออกแบบชุดเมทาดาทาสำหรับอธิบายเลิร์นนิงอ็อบเจกต์

ชุดเมทาดาทาที่ใช้อธิบายเลิร์นนิงอ็อบเจกต์ในระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ในงานวิจัยนี้ ออกแบบและพัฒนาขึ้นจากข้อมูลเมทาดาทา 2 ส่วน ได้แก่

1) เมทาดาทาที่ได้จากการศึกษาพฤติกรรมค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน พบเมทาดาทาที่ผู้ใช้เลือกใช้ในการพิจารณารายละเอียดของเลิร์นนิงอ็อบเจกต์ในระดับ “มาก” ขึ้นไป รวมจำนวน 10 องค์ประกอบ ได้แก่ ชื่อเรื่อง (Title) คำสำคัญ (Keyword) ประเภทของไฟล์ (Format) รายละเอียด (Description) บทคัดย่อ (Abstract) หัวเรื่อง (Subject) ภาษา (Language) แหล่งที่มาของข้อมูล (Source) ระดับความยากง่ายของเนื้อหา (Difficulty) ผู้รับผิดชอบเนื้อหา (Publisher)

2) เมทาดาทาที่ศึกษาจากงานวิจัยและข้อกำหนดตามมาตรฐานโครงสร้างด้านกายภาพและเนื้อหาของเลิร์นนิงอ็อบเจกต์ที่ควรมี จะได้องค์ประกอบเมทาดาทาเพิ่มเติมในชุดโครงสร้างอีก 7 องค์ประกอบ ได้แก่ ระดับการศึกษาหรือกลุ่มเป้าหมาย (EducationLevel/Audience) รูปแบบการปฏิสัมพันธ์ (Interactive Type) ระดับการปฏิสัมพันธ์ (Interactive Level) ประเภทเลิร์นนิงอ็อบเจกต์ (Learning object Type) ระยะเวลาการเรียน (Learning Time) มีส่วนประกอบ (hasPart) เป็นส่วนประกอบของ (isPartOf)

เมื่อรวบรวมเมทาดาทาจากทั้ง 2 ส่วนจะได้เป็นชุดเมทาดาทาที่ใช้อธิบายเลิร์นนิงอ็อบเจกต์ในงานวิจัยนี้ทั้งสิ้น 17 องค์ประกอบ โดยนำมาปรับเข้ากับเมทาดาทามาตรฐานสากล ที่นิยมใช้อธิบายเลิร์นนิงอ็อบเจกต์ได้แก่ มาตรฐาน Dublin Core มาตรฐาน และ IEEE LOM มาวิเคราะห์และกำหนดชุดขององค์ประกอบเมทาดาทาที่สามารถบรรยายลักษณะทางบรรณานุกรมเพื่อการสืบค้นเลิร์นนิงอ็อบเจกต์ ตามตารางที่ 4.14

ตารางที่ 4.14 เมทาดาทาที่ใช้ในการอธิบายเลิร์นนิงอ็อบเจกต์

องค์ประกอบ	มาตรฐาน	คำอธิบาย	ช่วงข้อมูล
title	Dublin Core	ชื่อบทเรียน	ข้อความชื่อทรัพยากร
subject	Dublin Core	หัวเรื่อง	หมวดหมู่เอกสารหรือคำแทนเอกสาร

ตารางที่ 4.14 เมทาดาทาที่ใช้ในการอธิบายเลิร์นนิงอ็อบเจกต์ (ต่อ)

องค์ประกอบ	มาตรฐาน	คำอธิบาย	ช่วงข้อมูล
abstract	Dublin Core	บทคัดย่อ	ข้อความอธิบายเนื้อหาย่อ
format	Dublin Core	รูปแบบสื่อ	ภาพเคลื่อนไหว ภาพนิ่ง เว็บไซต์ ไฟล์เสียง ข้อความตัวอักษร
description	Dublin Core	คำอธิบาย	วัตถุประสงค์ คำอธิบายเนื้อหา
language	Dublin Core	ภาษาที่ใช้	ภาษาไทย ภาษาอังกฤษ ภาษาอื่น ๆ
publisher	Dublin Core	สำนักพิมพ์ ผู้เผยแพร่	ชื่อบุคคล หน่วยงาน สำนักพิมพ์ ผู้รับผิดชอบการนำเสนอ
source	Dublin Core	แหล่งข้อมูล	แหล่งข้อมูลต้นฉบับ หรือ ได้รับข้อมูลมา
Educational.difficulty	LOM	ระดับความยาก	ง่ายมาก ง่าย ปานกลาง ยาก ยากมาก
General.keyword	LOM	คำสำคัญ	คำสำคัญในบทเรียน
Educational.InteractiveLevel	LOM	ระดับการมี ปฏิสัมพันธ์	น้อย ปานกลาง มาก
Educational.Typical LearningTime		ระยะเวลาในการเรียน	ระยะเวลาในการเรียน เช่น 30 นาที เป็นต้น

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตารางที่ 4.14 เมทาดาทาที่ใช้ในการอธิบายเลิร์นนิงอ็อบเจกต์ (ต่อ)

องค์ประกอบ	มาตรฐาน	คำอธิบาย	ช่วงข้อมูล
Educational.InteractiveType	LOM	ประเภทการมีปฏิสัมพันธ์	อ่าน ฟัง ดู กดปุ่ม พิมพ์ ตอบโต้หลายรูปแบบ
Educational. LearningObjectResourceType	LOM	ประเภทสื่อ	แบบฝึกหัด แบบจำลองเหตุการณ์ แบบทดสอบ การทดลอง โจทย์ปัญหา บทบรรยายเนื้อหา แบบสอบถาม แผนภาพ รูปภาพ กราฟ ตาราง สไลด์นำเสนอ แบบการเรียนรู้ด้วยตนเอง
EducationLevel	LOM	ระดับการศึกษา	มัธยมศึกษาชั้นปีที่ 1 มัธยมศึกษาชั้นปีที่ 2 มัธยมศึกษาชั้นปีที่ 3 มัธยมศึกษาชั้นปีที่ 4 มัธยมศึกษาชั้นปีที่ 5 มัธยมศึกษาชั้นปีที่ 6
hasPart	dcterms	มีส่วนประกอบ	วัตถุองค์ประกอบย่อย
isPartOf		เป็นวนประกอบของ	วัตถุองค์ประกอบหลัก

จากตารางที่ 4.14 รายการเมทาดาทาข้างต้นจะไปใช้ในการออกแบบฐานข้อมูลเชิงสัมพันธ์และพัฒนาเป็นระบบต้นแบบสำหรับจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ในลำดับถัดไป

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 5
การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา

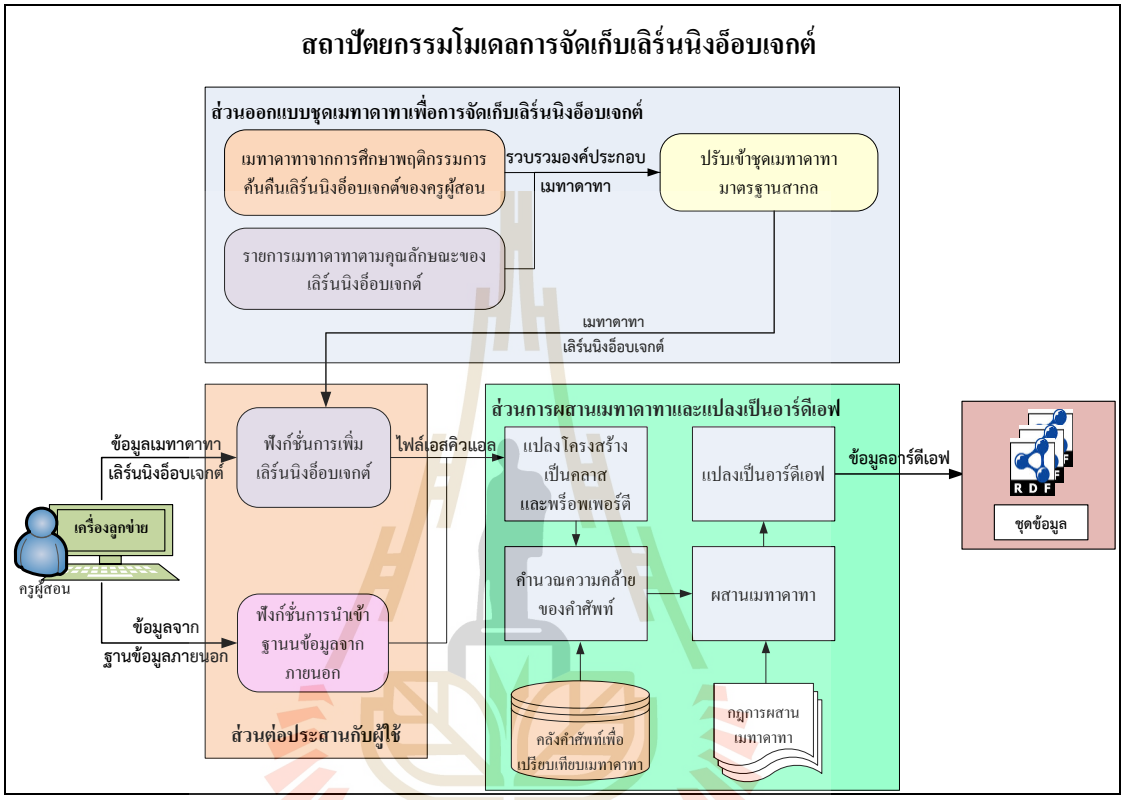
บทนี้นำเสนอผลการออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากข้อมูลโอเพนดาทา และการพัฒนาระบบต้นแบบเพื่อทดสอบการทำงานตามกระบวนการของโมเดลที่ออกแบบไว้ เพื่อแสดงให้เห็นว่าโมเดลสามารถรองรับการจัดเก็บเลิร์นนิงอ็อบเจ็กต์ โดยผู้ใช้ทำการเพิ่มข้อมูลเข้าสู่ระบบได้โดยตรงผ่านส่วนต่อประสานกับผู้ใช้ทางเว็บไซต์ต้นแบบ และสามารถนำเข้าข้อมูลที่มีอยู่เดิมจากระบบฐานข้อมูลที่มีเค้าร่างเมทาดาทาต่างกันได้ผ่านกระบวนการผสานเมทาดาทา ก่อนจะแปลงให้อยู่ในรูปแบบโครงสร้างอาร์ดีเอฟ (RDF) และยังคงความหมายและรูปแบบความสัมพันธ์ของข้อมูลไว้ได้อย่างถูกต้อง นอกจากกระบวนการจัดเก็บข้อมูลในรูปแบบโอเพนดาทาแล้ว โมเดลนี้ยังสามารถค้นคืนข้อมูล เลิร์นนิงอ็อบเจกต์ที่มีความสอดคล้องและสัมพันธ์กับคำสำคัญที่ใช้สืบค้น โดยใช้เทคนิคการวัดความคล้ายคลึงเชิงความหมาย ผ่านคลังคำศัพท์เพื่อเชื่อมโยงกับหัวเรื่องที่เกี่ยวข้องรวมถึงสามารถจัดเรียงผลการสืบค้นตามเงื่อนไขของครูผู้สอนได้ โดยในงานวิจัยนี้ใช้เลิร์นนิงอ็อบเจกต์รายวิชาคณิตศาสตร์ในระดับมัธยมศึกษาชั้นปีที่ 1-6 เป็นข้อมูลในการพัฒนาระบบต้นแบบ

เนื้อหาภายในบทนี้ประกอบด้วย (1) ผลการออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ (2) ผลการออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ และ 3) ผลการประเมินการทำงานของระบบการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์แบ่งการประเมินเป็น 2 รูปแบบ ได้แก่การประเมินประสิทธิภาพการทำงานระบบ และการประเมินการทำงานโดยผู้ใช้งาน ซึ่งมีรายละเอียดดังต่อไปนี้

5.1 ผลการออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์

โมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์มีกระบวนการทั้งหมด 2 ส่วนได้แก่ (1) การออกแบบชุดเมทาดาทาเพื่อการจัดเก็บเลิร์นนิงอ็อบเจกต์ โดยออกแบบชุดเมทาดาทาจากการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน และการศึกษาองค์ประกอบของเลิร์นนิงอ็อบเจกต์ที่จำเป็นสำหรับการแลกเปลี่ยนข้อมูลและการใช้ข้อมูลซ้ำ โดยนำมาปรับเข้ากับเมทาดาทามาตรฐานสากลและนำไปพัฒนาเป็นส่วนต่อประสานกับผู้ใช้ผ่านเว็บไซต์ (2) การผสานเมทาดาทาของเลิร์นนิงอ็อบเจกต์ที่มีโครงสร้าง

เมทาดาทาที่แตกต่างกันและแปลงโครงสร้างเป็นอาร์ดีเอฟ เพื่อรองรับการนำเข้าข้อมูลจากภายนอกระบบ พร้อมทั้งทำการแปลงโครงสร้างให้เป็นชุดข้อมูลอาร์ดีเอฟ (RDF) เพื่อรองรับมาตรฐานโอเพนดาทาสำหรับการแลกเปลี่ยนข้อมูลแบบเปิดเสรี โดยแสดงสถาปัตยกรรมโมเดลดังรูปที่ 5.1



รูปที่ 5.1 สถาปัตยกรรมโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์

5.1.1 การออกแบบเมทาดาทาเพื่อการจัดเก็บเลิร์นนิงอ็อบเจกต์

เมทาดาทาสำหรับบรรยายคุณลักษณะของเลิร์นนิงอ็อบเจกต์ที่จัดเก็บในคลังข้อมูลเลิร์นนิงอ็อบเจกต์ของงานวิจัยนี้ ใช้ข้อมูลจากการศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอนร่วมกับเมทาดาทาตามลักษณะของเลิร์นนิงอ็อบเจกต์สำหรับการแลกเปลี่ยนข้อมูลและการใช้ซ้ำ โดยใช้มาตรฐานดับลินคอร์ (Dublin Core) และมาตรฐานลอม (LOM) มาเป็นเมทาดาทามาตรฐานในการอธิบายข้อมูลของเลิร์นนิงอ็อบเจกต์เพื่อนำมาออกแบบระบบฐานข้อมูล โดยใช้ตัวอย่างข้อมูลเลิร์นนิงอ็อบเจกต์จากเว็บไซต์ที่เผยแพร่เลิร์นนิงอ็อบเจกต์ในรายวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาชั้นปีที่ 1-6 ซึ่งเป็นข้อมูลแบบเปิดเสรี โดยแบ่งออกเป็น 10 คลาส ได้แก่ ผู้สร้างผลงาน (Contributor) แผนกและส่วนงาน (Department) หน่วยบทเรียน (Lesson) วัตถุประยุกต์ (Application Object) วัตถุสารสนเทศ (Information

Object) สื่อข้อมูลย่อย (Asset) คลาสคำสำคัญ (Keyword) คลาสรูปแบบการมีปฏิสัมพันธ์ (InteractiveType) คลาสประเภทเลิร์นนิงอ็อบเจกต์ (ResourceType) และคลาสภาษา (Languages) โดยมีรายละเอียดดังนี้

ตารางที่ 5.1 คลาสผู้สร้างผลงาน (Contributor)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
ContributeId	string	รหัสผู้สร้างผลงาน
ConName	string	ชื่อผู้สร้างผลงาน
ConSurName	string	ชื่อสกุลผู้สร้างผลงาน
Position	string	ตำแหน่ง
Role	string	บทบาทที่รับผิดชอบ
Email	string	อีเมล
Address	string	ที่อยู่
Province	string	จังหวัด
Zipcode	string	รหัสไปรษณีย์

คลาสผู้สร้างผลงานเป็นคลาสที่เก็บข้อมูลของครูผู้สอนหรือผู้ที่มีส่วนร่วมในการสร้างชิ้นงานโดยครอบคลุมหน้าที่ทั้งผู้แต่งเนื้อหา ผู้วาดภาพประกอบ ผู้ตรวจสอบพิสูจน์อักษร และหน้าที่อื่น ๆ ที่ส่วนในการทำให้ผลงานเสร็จสมบูรณ์ โดยในงานวิจัยนี้ ผู้สร้างผลงานจะไม่เกี่ยวข้องกับผู้เพิ่มข้อมูลหรือกรอกข้อมูลลงในระบบ

ตารางที่ 5.2 คลาสแผนกและส่วนงาน (Department)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
DeptId	string	รหัสแผนก
DeptName	string	ชื่อแผนก
InstituteName	string	สถาบัน
Description	string	รายละเอียดหน่วยงาน
Address	string	ที่อยู่หน่วยงาน

คลาสแผนกและส่วนงานสำหรับเก็บข้อมูลของส่วนงาน เพื่อใช้อธิบายความสัมพันธ์กับผู้สร้างผลงานว่าสังกัดอยู่ส่วนงานใด ภายใต้สังกัดใด เพื่อใช้เป็นข้อมูลในการติดต่อประสานงาน

คลาสหน่วยบทเรียน (Lesson) คลาสวัตถุประยุกต์ (Application Object) และคลาสวัตถุสารสนเทศ (Information Object) เป็นคลาสที่ใช้เก็บข้อมูลเลิร์นนิงอ็อบเจกต์โดยแบ่งตามองค์ประกอบเนื้อหาของเลิร์นนิงอ็อบเจกต์แต่ละระดับ ดังนั้นจะมีชุดเมทาดาทาหลักเหมือนกัน

ตารางที่ 5.3 คลาสหน่วยบทเรียน (Lesson)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
lessonId	string	รหัสบทเรียน
Title	string	ชื่อบทเรียน
Abstract	string	บทคัดย่อของเนื้อหา
Description	string	คำอธิบายเกี่ยวกับบทเรียน
Objective	string	วัตถุประสงค์บทเรียน
EducationLevel	string	ระดับการศึกษา
Difficulty	string	ระดับความยาก
InteractiveLevel	string	ระดับการมีส่วนร่วม
LearningTime	integer	ระยะเวลาในการเรียน
Publisher	string	สำนักพิมพ์หรือ ผู้เผยแพร่
Source	string	แหล่งข้อมูลต้นฉบับ
DateCreated	string	วันที่สร้างบทเรียนขึ้น
LastModified	string	วันที่ปรับปรุงข้อมูลล่าสุด
hasPart	string	มีส่วนประกอบเลิร์นนิงอ็อบเจกต์อื่นในบทเรียน

คลาสหน่วยบทเรียนเป็นคลาสจัดเก็บข้อมูลรายละเอียดของบทเรียน โดยหน่วยบทเรียนประกอบด้วย รหัสบทเรียน ชื่อบทเรียน บทคัดย่อ คำอธิบายบทเรียน วัตถุประสงค์บทเรียน ระดับการศึกษา ระดับความยาก ระดับการมีส่วนร่วมในเลิร์นนิงอ็อบเจกต์ เวลาที่ใช้ในการเรียน สำนักพิมพ์หรือผู้เผยแพร่ แหล่งที่มาของข้อมูล วันที่สร้างและปรับปรุงล่าสุด และข้อมูลส่วนประกอบของเลิร์นนิงอ็อบเจกต์อื่นในบทเรียน

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตารางที่ 5.4 คลาสวัตถุประยุกต์ (ApplicationObject)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
AOId	string	รหัสวัตถุประยุกต์
Title	string	ชื่อเนื้อหา
Abstract	string	บทคัดย่อของเนื้อหา
Description	string	คำอธิบายเกี่ยวกับเนื้อหา
Objective	string	วัตถุประสงค์เนื้อหา
EducationLevel	string	ระดับการศึกษา
Difficulty	string	ระดับความยาก
InteractiveLevel	string	ระดับการมีส่วนร่วม
LearningTime	integer	ระยะเวลาในการเรียน
Publisher	string	สำนักพิมพ์หรือ ผู้เผยแพร่
Source	string	แหล่งข้อมูลต้นฉบับ
DateCreated	string	วันที่สร้างบทเรียนขึ้น
LastModified	string	วันที่ปรับปรุงข้อมูลล่าสุด
hasPart	string	มีส่วนประกอบเลิร์นนิงอ็อบเจกต์อื่นในบทเรียน
isPartOf	string	เป็นส่วนประกอบของเลิร์นนิงอ็อบเจกต์รายการอื่น

ตารางที่ 5.5 คลาสวัตถุสารสนเทศ (InformationObject)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
IOId	string	รหัสวัตถุสารสนเทศ
Title	string	ชื่อเนื้อหา
Abstract	string	บทคัดย่อของเนื้อหา
Description	string	คำอธิบายเกี่ยวกับเนื้อหา
Objective	string	วัตถุประสงค์เนื้อหา
EducationLevel	string	ระดับการศึกษา
Difficulty	string	ระดับความยาก

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตารางที่ 5.5 คลาสวัตถุสารสนเทศ (InformationObject) (ต่อ)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
InteractiveLevel	string	ระดับการมีส่วนร่วม
LearningTime	integer	ระยะเวลาในการเรียน
Publisher	string	สำนักพิมพ์หรือ ผู้เผยแพร่
Source	string	แหล่งข้อมูลต้นฉบับ
DateCreated	string	วันที่สร้างบทเรียนขึ้น
LastModified	string	วันที่ปรับปรุงข้อมูลล่าสุด
hasPart	string	มีส่วนประกอบเลิร์นนิง อ็อบเจกต์อื่นในบทเรียน
isPartOf	string	เป็นส่วนประกอบของ เลิร์นนิงอ็อบเจกต์รายการอื่น

ตารางที่ 5.6 คลาสรายการข้อมูลย่อย (Asset)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
AssetID	string	รหัสชิ้นงาน
AssetTitle	string	ชื่อชิ้นงาน
AssetType	string	ประเภทชิ้นงาน
Format	string	รูปแบบชิ้นงาน
Encoding	string	การเข้ารหัส
Dimenstion	string	ข้อมูลทางกายภาพ
DownloadCount	integer	จำนวนการดาวน์โหลด
FileName	string	ชื่อไฟล์
PathStorage	string	ที่จัดเก็บไฟล์
Status	string	สถานะไฟล์
isPartOf	string	เป็นส่วนประกอบของ เลิร์นนิงอ็อบเจกต์รายการอื่น

มหาวิทยาลัยเทคโนโลยีสุรนารี

คลาสรายการข้อมูลย่อย (Asset) อธิบายถึงตำแหน่งและการจัดเก็บไฟล์ข้อมูล เก็บข้อมูลรหัสชิ้นงาน ชื่อชิ้นงาน ประเภทของชิ้นงาน รูปแบบไฟล์ การเข้ารหัส ข้อมูลทางกายภาพ จำนวนการดาวน์โหลด ที่จัดเก็บไฟล์และสถานะของไฟล์

ตารางที่ 5.7 คลาสคำสำคัญ (Keyword)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
KeywordID	string	รหัสคำสำคัญ
Keyword	string	คำสำคัญ

ตารางที่ 5.8 คลาสรูปแบบการมีปฏิสัมพันธ์ (InteractiveType)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
ITTypeId	string	รหัสรูปแบบปฏิสัมพันธ์
ITType	string	รูปแบบปฏิสัมพันธ์

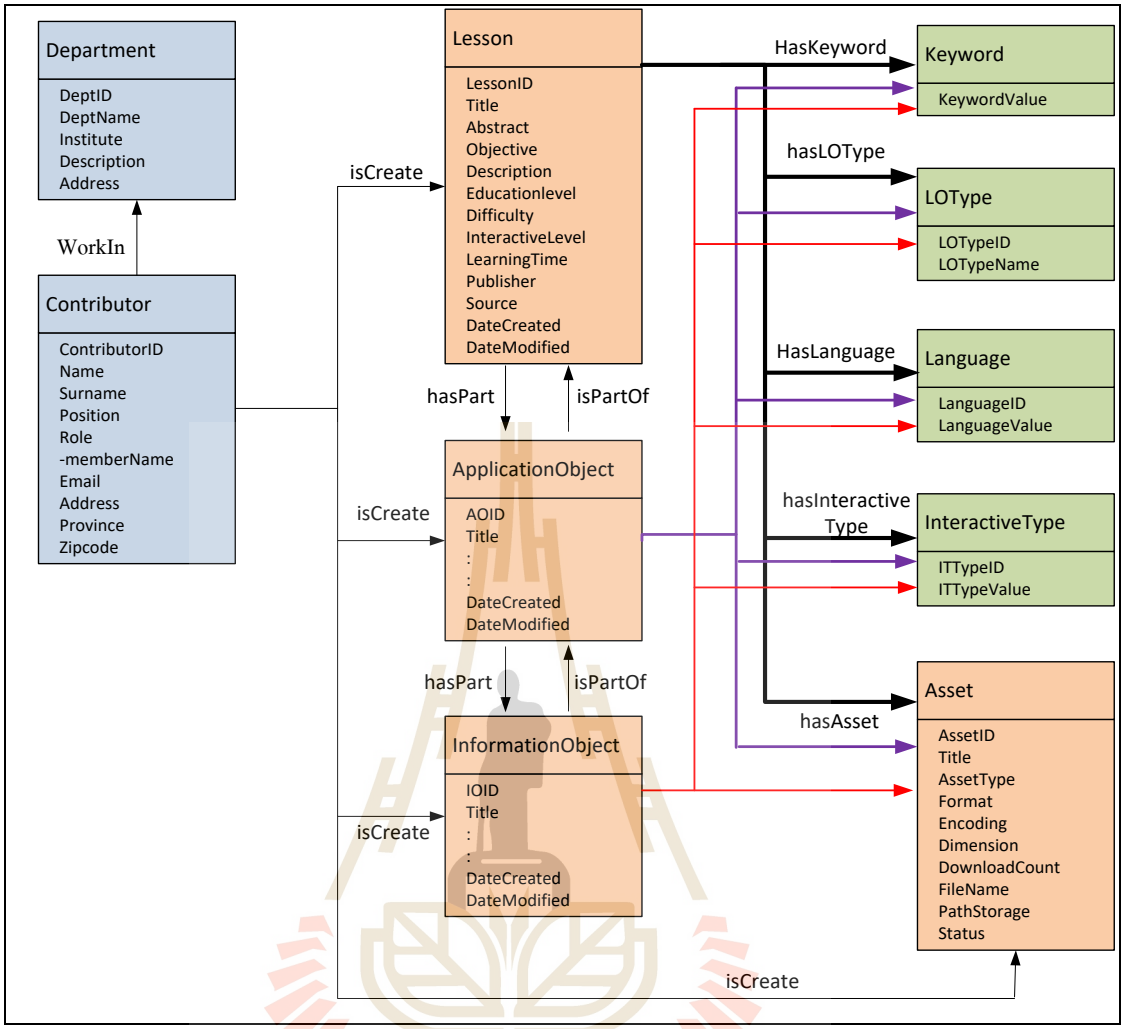
ตารางที่ 5.9 คลาสประเภทเลิร์นนิงอ็อบเจกต์ (LOType)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
LOTypeId	string	รหัสประเภท
LOType	string	ประเภทสื่อ

ตารางที่ 5.10 คลาสภาษา (Language)

คุณลักษณะ	ชนิดข้อมูล	ความหมาย
LanguageId	string	รหัสภาษา
Language	string	ภาษา

คลาสคำสำคัญ (Keyword) คลาสรูปแบบการมีปฏิสัมพันธ์ (InteractiveType) คลาสประเภทเลิร์นนิงอ็อบเจกต์ (LOType) และคลาสภาษา (Language) เป็นคลาสสำหรับเก็บข้อมูลของเลิร์นนิงอ็อบเจกต์ที่มีค่าได้มากกว่าหนึ่งค่าในแต่ละระเบียน เช่น เลิร์นนิงอ็อบเจกต์หนึ่งระเบียนอาจจะมีการใช้ภาษาได้หลายภาษา หรือมีรูปแบบการปฏิสัมพันธ์ได้มากกว่าหนึ่งรูปแบบในเวลาเดียวกัน เป็นต้น



รูปที่ 5.2 โครงสร้างคลังข้อมูลสำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์

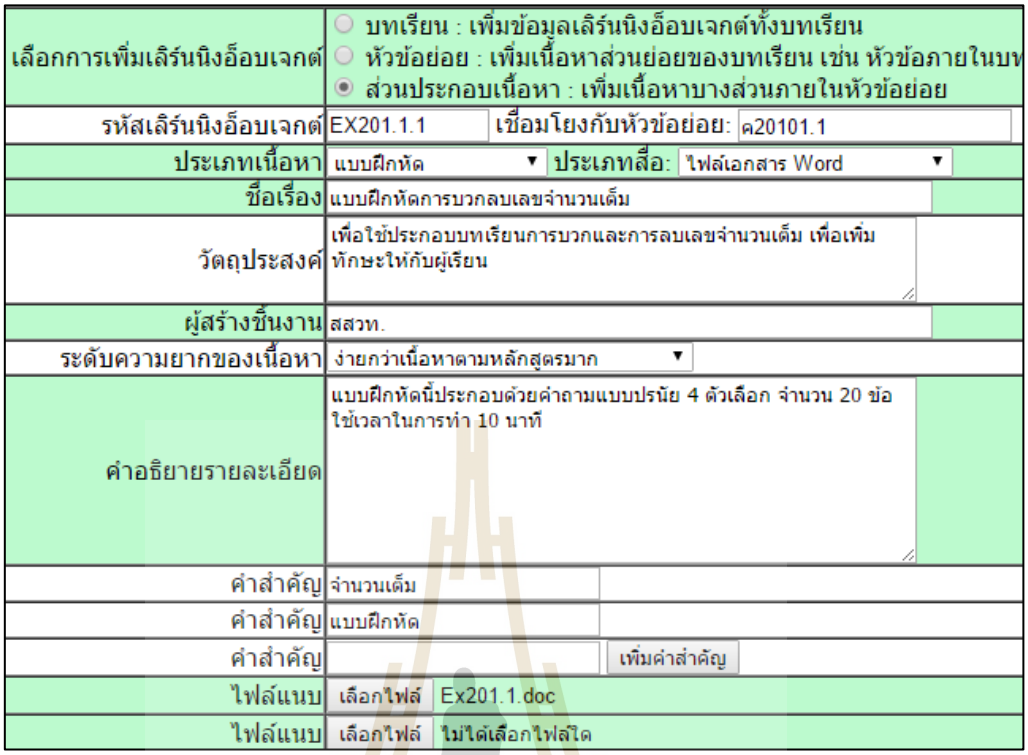
จากรูปที่ 5.2 แสดงโครงสร้างคลังข้อมูลที่ใช้ในการจัดเก็บเลิร์นนิงอ็อบเจกต์ที่สามารถแยกระดับออกเป็น 3 ระดับได้แก่ หน่วยบทเรียน (Lesson) วัตถุประยุตก์ (Application Object) และวัตถุสารสนเทศ (Information Object) โดยเลิร์นนิงอ็อบเจกต์ทั้ง 3 ระดับจะมีความสัมพันธ์กันคือ หน่วยบทเรียนจะประกอบด้วยวัตถุประยุกต์หรือหัวข้อย่อยในบทเรียนได้หลายรายการซึ่งจะเรียบเรียงกันเป็นเนื้อหาตามลำดับการนำเสนอที่กำหนดไว้ ส่วนวัตถุประยุกต์จะมีส่วนประกอบเป็นวัตถุสารสนเทศซึ่งเป็นเลิร์นนิงอ็อบเจกต์หลาย ๆ ประเภท เช่น เนื้อหาบทเรียน แบบฝึกหัด กิจกรรมท้ายบท หรือข้อสอบ เป็นต้น ที่ประกอบอยู่ในวัตถุประยุกต์แต่ละรายการ ซึ่งเลิร์นนิงอ็อบเจกต์ทั้ง 3 ระดับจะสัมพันธ์กันผ่านพร็อพเพอร์ตี hasPart และ isPartOf เพื่อเชื่อมโยงความสัมพันธ์ระหว่างกัน และมีความสัมพันธ์กับคลาสคำสำคัญ (Keyword) ผ่าน hasKeyword

สำหรับเชื่อมโยงคำดัชนีในการเข้าถึงเลิร์นนิงอ็อบเจกต์ และสัมพันธ์กับคลาสภาษา (Language) ด้วย hasLanguage เพื่ออธิบายข้อมูลภาษาภายในเนื้อหาของเลิร์นนิงอ็อบเจกต์แต่ละรายการสามารถมีได้หลายภาษา เช่นเดียวกับคลาสประเภทของเลิร์นนิงอ็อบเจกต์ (LOType) ที่เลิร์นนิงอ็อบเจกต์แต่ละรายการจะประกอบด้วยเลิร์นนิงอ็อบเจกต์ระดับเล็กกว่าหลาย ๆ ประเภทมาร่วมกัน เช่น ในหนึ่งหัวข้อย่อย จะประกอบด้วยเนื้อหาบทเรียน แบบฝึกหัด และกิจกรรมท้ายบท หรือเกมเสริมความรู้ เป็นต้น โดยจะสัมพันธ์ผ่านความสัมพันธ์ hasLOType และมีความสัมพันธ์กับคลาสรูปแบบการมีปฏิสัมพันธ์ (InteractiveType) ผ่าน hasITType เพื่ออธิบายเลิร์นนิงอ็อบเจกต์แต่ละรายการว่าสามารถมีการปฏิสัมพันธ์รูปแบบใด เช่น การดู อ่าน พูด เขียน หรือตอบโต้ด้วยวิธีอื่น ๆ ตามที่บทเรียนกำหนด เป็นต้น เลิร์นนิงอ็อบเจกต์แต่ละระดับจะมีองค์ประกอบระดับย่อยที่สุดคือสื่อข้อมูลย่อย (Asset) ที่ใช้เป็นเสมือนวัตถุดิบในการสร้างเลิร์นนิงอ็อบเจกต์แต่ละรายการขึ้นมาผ่านความสัมพันธ์ hasAsset ซึ่งข้อมูลย่อยแต่ละรายการมีคุณสมบัติในการใช้ซ้ำ โดยสามารถนำไปประกอบเป็นเลิร์นนิงอ็อบเจกต์อื่น ๆ ได้

นอกจากข้อมูลเลิร์นนิงอ็อบเจกต์ระบบยังรองรับการเก็บข้อมูลของคลาสผู้สร้างผลงาน (Contributor) โดยมีการเชื่อมโยงข้อมูลกับหน่วยงานต้นสังกัด (Department) ผ่านความสัมพันธ์ WorkIn เพื่ออธิบายรายละเอียดของต้นสังกัดและบทบาทหน้าที่ในการจัดทำเลิร์นนิงอ็อบเจกต์ในแต่ละระดับ

จากรายละเอียดของเมทาดาทาที่ออกแบบไว้ได้นำไปพัฒนาเป็นระบบฐานข้อมูลสำหรับจัดเก็บเลิร์นนิงอ็อบเจกต์และพัฒนาส่วนต่อประสานกับผู้ใช้บนเว็บไซต์ระบบต้นแบบ โดยแยกเป็นฟังก์ชั่นการเพิ่มข้อมูลเลิร์นนิงอ็อบเจกต์ทีละรายการ และฟังก์ชั่นการนำเข้าฐานข้อมูลเลิร์นนิงอ็อบเจกต์จากภายนอก โดยได้ดังภาพที่ 5.3 เป็นหน้าจอส่วนต่อประสานกับผู้ใช้ส่วนการเพิ่มข้อมูลเลิร์นนิงอ็อบเจกต์รายการใหม่เข้าสู่ระบบโดยตรงผ่านฟอร์มกรอกข้อมูล

มหาวิทยาลัยเทคโนโลยีสุรนารี

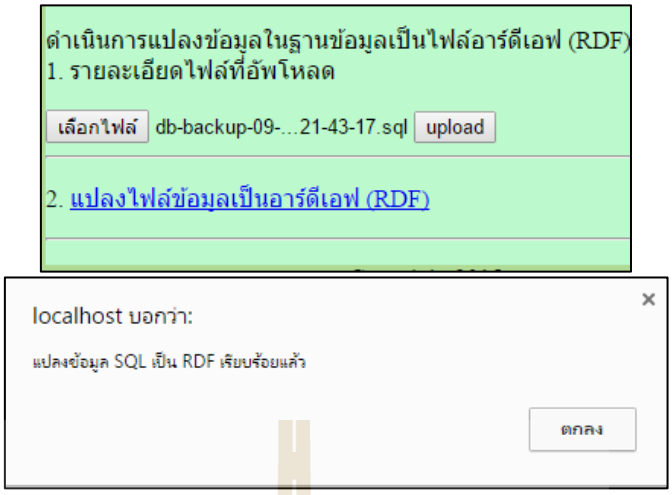


รูปที่ 5.3 ฟังก์ชั่นการเพิ่มข้อมูลเลิร์นนิงอ็อบเจกต์

จากรูปที่ 5.3 เป็นหน้าจอสำหรับการเพิ่มเลิร์นนิงอ็อบเจกต์เข้าสู่ฐานข้อมูล การเพิ่มเลิร์นนิงอ็อบเจกต์จะให้ผู้ใช้เลือกลักษณะของเลิร์นนิงอ็อบเจกต์ใน 3 รูปแบบได้แก่ เลิร์นนิงอ็อบเจกต์ระดับบทเรียน เลิร์นนิงอ็อบเจกต์ระดับหัวข้อย่อย และเลิร์นนิงอ็อบเจกต์ระดับส่วนประกอบเนื้อหา ซึ่งผู้ใช้สามารถทำการเพิ่มข้อมูลตามเมทาดาทาที่กำหนดไว้และเพิ่มไฟล์แนบในตอนท้ายของแบบฟอร์มเพื่อนำไปจัดเก็บในฐานข้อมูล

5.1.2 การผสานเมทาดาทาและการแปลงโครงสร้างอาร์ดีเอฟ

ฟังก์ชั่นการผสานเมทาดาทาและแปลงโครงสร้างอาร์ดีเอฟจะทำงานแบบอัตโนมัติเมื่อผู้ใช้นำเข้าระบบฐานข้อมูลที่เป็นไฟล์เอสคิวแอล (SQL) จากฐานข้อมูลเชิงสัมพันธ์ และเลือกอัพโหลดเข้าสู่กระบวนการแปลงโครงสร้างข้อมูลเป็นอาร์ดีเอฟ จากนั้นระบบจะทำการแปลงข้อมูลเป็นไฟล์อาร์ดีเอฟพร้อมกับผสานเมทาดาทาตามกฎที่กำหนดไว้เบื้องหลัง ดังรูปที่ 5.4



รูปที่ 5.4 ฟังก์ชั่นการผสานเมทาดาทาและการแปลงโครงสร้างอาร์ดีเอฟ

จากรูปที่ 5.4 ฟังก์ชั่นการนำเข้าฐานข้อมูลจากภายนอกที่สำรองข้อมูลมาเป็นไฟล์ข้อมูลเอสคิวแอล เมื่อเลือกไฟล์ที่ต้องการนำเข้าแล้วระบบจะดำเนินการผสานเมทาดาทาและแปลงโครงสร้างเป็นอาร์ดีเอฟ โดยผ่านกระบวนการ 4 ขั้นตอน ดังนี้

(1) การปรับโครงสร้างฐานข้อมูลเป็นคลาสและพร็อพเพอร์ตี

ในขั้นตอนนี้ข้อมูลที่นำเข้าสู่กระบวนการผสานเมทาดาทาจะถูกปรับเปลี่ยนโครงสร้างจากข้อมูลที่เป็นความสัมพันธ์แบบตารางข้อมูลเชิงสัมพันธ์ให้อยู่ในรูปแบบของคลาส (Class) และคุณสมบัติของคลาส (Data Properties)โดยการดึงชื่อตารางแต่ละตารางพร้อมทั้งนำข้อมูลแอตทริบิวต์ภายในตารางมาการปรับให้เป็นคลาสและคุณสมบัติของคลาส และในส่วนความสัมพันธ์ระหว่างตารางก็จะมีการเปลี่ยนรูปแบบให้เป็นความสัมพันธ์ระหว่างคลาส (Object Properties) โดยขั้นตอนสามารถแสดงได้ดังรูปที่ 5.5 เป็นคำสั่งการผสานหรือแปลงตารางข้อมูลให้เป็นคลาส และแปลงแอตทริบิวต์ให้เป็นพร็อพเพอร์ตี ซึ่งกฎการแปลงตารางข้อมูลเป็นคลาสนี้จะกำหนดด้วยคำศัพท์สัมพันธ์ของเมทาดาทาทางคณิตศาสตร์ โดยทำงานร่วมกับชุดคำสั่งด้วย D2RQ API ซึ่งเป็นชุดคำสั่งเปิดเผยรหัสทำหน้าที่ช่วยในการจัดเก็บโครงสร้างฐานข้อมูลและข้อมูลภายในตารางข้อมูลก่อนทำการแปลงเป็นโครงสร้างอาร์ดีเอฟ ซึ่งเป็นตัวอย่างการแปลงตารางของข้อมูลย่อย (Asset Table) ให้เป็นคลาส ด้วยคำสั่ง ClassMap และแปลงแอตทริบิวต์รหัสข้อมูลย่อย (AssetId) เป็น “พร็อพเพอร์ตี” AssetId ด้วยคำสั่ง PropertiesBridge โดยกระบวนการนี้จะทำการแปลงตารางข้อมูลทั้งหมดภายในฐานข้อมูลเก็บไว้ในหน่วยความจำเพื่อทำการคำนวณค่าความคล้ายคลึงของคำศัพท์ต่อไป

```
map:asset a d2rq:ClassMap;
        d2rq:dataStorage map:database;
        d2rq:uriPattern "asset/@@asset.AssetId|urlify@@";
        d2rq:class vocab:asset;
        d2rq:classDefinitionLabel "asset";
        .
map:asset__label a d2rq:PropertyBridge;
        d2rq:belongsToClassMap map:asset;
        d2rq:property rdfs:label;
        d2rq:pattern "asset #@@asset.AssetId@@";
        .
map:asset_AssetId a d2rq:PropertyBridge;
        d2rq:belongsToClassMap map:asset;
        d2rq:property vocab:asset_AssetId;
        d2rq:propertyDefinitionLabel "asset AssetId";
        d2rq:column "asset.AssetId";
```

รูปที่ 5.5 คำสั่งแปลงตารางข้อมูลให้เป็นคลาสและพร็อพเพอร์ตี

(2) การคำนวณค่าความคล้ายคลึงของคำศัพท์ระหว่างชุดข้อมูล

การระบวนการคำนวณค่าความคล้ายคลึงของคำศัพท์ (ชื่ออิลิเมนต์กับชื่อฟิลด์) จะทำหลังจากปรับโครงสร้างของฐานข้อมูลให้อยู่ในรูปของคลาสและพร็อพเพอร์ตีแล้ว จากนั้นข้อมูลแต่ละรายการจะถูกนำมาเข้ากระบวนการวัดค่าความคล้ายคลึง โดยมีชุดของเมทาดาทาที่สร้างขึ้นเป็นต้นแบบสำหรับการเปรียบเทียบ โดยกระบวนการเปรียบเทียบและผสานเมทาดาทาจะใช้แนวคิดการวัดความคล้ายคลึงด้วยแนวคิดของแจ็คการ์ด โดยการวัดความคล้ายคลึงจะได้ผลคะแนนความคล้ายคลึงของคำศัพท์จากความสัมพันธ์ของเมทาดาทา ร่วมกับการพิจารณาลักษณะข้อมูลที่จัดเก็บที่มีลักษณะที่เหมือนกัน โดยมีรายละเอียดดังนี้

การคำนวณหาค่าความคล้ายคลึงระหว่างเอนทิตีหรือคลาสของแหล่งข้อมูลต้นทาง (Source) และเอนทิตีหรือคลาสเป้าหมาย (Target) โดยใช้การพิจารณาตามโครงสร้างความสัมพันธ์ของคำศัพท์ โดยการนำเครื่องมือการวัดความคล้ายคลึงตามโครงสร้างของคำจากเวิร์ดเน็ต (WordNet Similarity Package) ร่วมกับสมการการวัดความคล้ายคลึงของวูและพาร์ลเมอร์ (Wu and Palmer, 1994) โดยพิจารณาความลึกของ 2 ชุดข้อมูล (Synset) ภายในโครงสร้างศัพท์สัมพันธ์ของเวิร์ดเน็ต ดังนี้

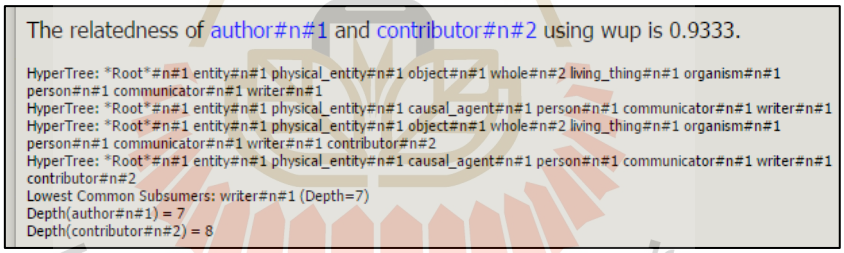
$$sim_{WuP}(c_1 c_2) = \frac{2*depth(LCS)}{(depth(c_1))+(depth(c_2))} \qquad (5\text{-}1)$$

จากสมการค่าของการคำนวณความคล้ายคลึงอยู่ระหว่าง 0 < ค่า Simwup <= 1 โดยค่าจากสมการจะมีค่ามากกว่าศูนย์เสมอ เนื่องจากค่าความลึกของ LCS ที่คำนวณได้ต้องไม่เท่ากับศูนย์ ทั้งนี้ค่าความคล้ายคลึงจะเท่ากับ 1 ได้นั้น เกิดจากกรณีที่คำทั้งสองอยู่ในซิงค์เซ็ตหรือกลุ่มคำศัพท์เดียวกัน ตัวอย่างการหาค่าความคล้ายคลึงตามสมการของวูและพาร์มเมอร์ (Wu and Palmer) ของคำว่า Author และ Contributor ซึ่งทั้งสองคำนี้ใช้สำหรับเป็นเมทาดาทาเก็บชื่อผู้สร้างผลงานหรือผู้เขียน ซึ่งสามารถใช้ได้ทั้งสองคำ เมื่อคำนวณความคล้ายคลึงตามโครงสร้างคำของเวิร์ดเน็ตแล้ว ผลลัพธ์ที่ได้คือ

$$sim_{WuP}(author, contributor) = \frac{2 * depth(LCS)}{(depth(author)) + (depth(contributor))}$$

$$sim_{WuP}(author, contributor) = \frac{2 * 7}{7 + 8}$$

$$sim_{WuP}(author, contributor) = 0.93$$



รูปที่ 5.6 การคำนวณค่าความคล้ายคลึงของคำศัพท์เมทาดาทา

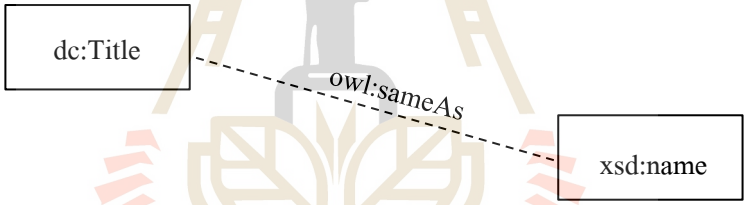
จากภาพที่ 5.6 เมื่อคำนวณตามสูตรของวูและพาร์เมอร์แล้วจะได้ค่าความคล้ายคลึงที่ 0.9333 ซึ่งหมายถึงคำสองคำนี้สามารถใช้แทนกันได้ หรือมีความหมายในหมวดหมู่คำศัพท์ที่ใกล้เคียงกันมาก ซึ่งจากผลการคำนวณสามารถแสดงได้ด้วยระยะทางระหว่างคำศัพท์ทั้งสองคือ LCS = 7 หมายถึงระยะห่างของคำทั้งสองในโครงสร้างคลังคำศัพท์จะห่างกันอยู่ 7 โหนด และระยะทางจากรากถึงโหนดคำ author และคำ contributor มีค่าเท่ากับ 7 และ 8 ตามลำดับ เมื่อใช้หลักการนี้พิจารณาคำทีละคู่จนครบทุกคำในชุดเมทาดาทาของทั้งสองชุด และจะทราบคะแนนความคล้ายคลึงเมื่อเปรียบเทียบระดับความลึกและชุดของคำศัพท์ที่ปรากฏอยู่ ซึ่งในงานวิจัยนี้กำหนดค่าความคล้ายคลึงตั้งแต่ 0.7 ขึ้นไป ซึ่งได้จากการทดสอบวัดความคล้ายคลึงคำศัพท์จาก

งานวิจัยอื่นถือว่าเป็นระดับความคล้ายคลึงที่เชื่อถือได้จากศัพท์สัมพันธ์ที่เวิร์ดเน็ต (Burdanitsky and Hirst: 2006) แต่ในบางกรณีโครงสร้างของเวิร์ดเน็ตก็มีผลคำนวณที่แตกต่างกันมาก แม้ว่าเมื่อพิจารณาถึงข้อมูลที่จัดเก็บเป็นข้อมูลประเภทหรือกลุ่มเดียวกันหรือมีความใกล้เคียงกัน จึงต้องมีการใช้แนวทางอื่นประกอบการพิจารณาความคล้ายคลึงนอกจากคำศัพท์ของเมทาดาทาแต่ละคำในลำดับถัดไป

(3) การผสานเมทาดาทา

(3.1) การพิจารณาและปรับปรุงเขตข้อมูลด้วยวิธีเชิงความหมาย

หลังจากการวัดความคล้ายคลึงของคลาสตามโครงสร้างคำศัพท์จากคลังคำศัพท์ของเมทาดาทาที่กำหนดแล้ว จะมีการเปรียบเทียบคลาสหรือพร็อพเพอร์ตีที่มีความคล้ายคลึงเชิงความหมายเพิ่มเติม โดยการกำหนดความสัมพันธ์ด้วยคุณสมบัติของอาว (OWL) เช่น owl:equivalentClass owl:equivalentProperty และ owl:sameAs เป็นต้น (Changqing, L., Ling, T. W., 2004; Seksun Suwanmanee, Djamal Benslimane and Philippe Thiran., 2005) เพื่อนำมาใช้ในการพิจารณาความสัมพันธ์ระหว่างคำศัพท์ที่ใช้ในเขตข้อมูล ดังรูปที่ 5.7



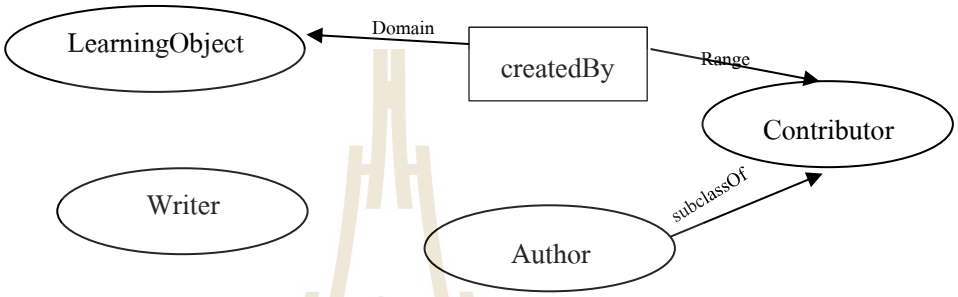
รูปที่ 5.7 การกำหนดคุณสมบัติระหว่างคำศัพท์ที่มีความคล้ายคลึงกัน

จากรูปที่ 5.7 การพิจารณาเขตข้อมูลภายในคลาสที่มีค่าความคล้ายคลึงกันด้วยคุณสมบัติ owl:sameAs เพื่อกำหนดว่าเขตข้อมูล Title และเขตข้อมูล xsd:name เหมือนกันหรือใช้แทนกันได้ ซึ่งวิธีการวิเคราะห์จะใช้วิธีการกำหนดชุดคำศัพท์ที่อยู่ในหมวดหมู่เดียวกันและสามารถใช้ร่วมกันหรือใช้แทนกันได้เป็นศัพท์สัมพันธ์ เพื่อนำมาวิเคราะห์ในการกำหนดความสัมพันธ์ข้างต้น ซึ่งชุดคลังคำนี้นำมาจากคำศัพท์ตามเค้าร่างเมทาดาทาของดับลินคอร์ และมาตรฐานลอมเป็นหลัก และเพิ่มคำศัพท์ตามคุณลักษณะที่ควรมีของเลิร์นนิงอ็อบเจกต์ เพื่อให้การเปรียบเทียบเขตข้อมูลมีความถูกต้องและครอบคลุมคำศัพท์ทีเกี่ยวข้องให้มากที่สุด

(3.2) การพิจารณารูปแบบความขัดแย้งของคำศัพท์ที่มีความคล้ายคลึงต่ำ

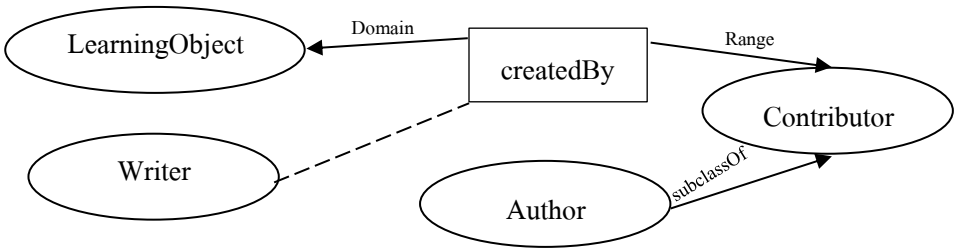
ในกระบวนการนี้ผู้วิจัยพิจารณาถึงโครงสร้างของคลาสทั้งสองโครงสร้างด้วยตนเองเพิ่มเติมจากการใช้การคำนวณความคล้ายคลึงเชิงความหมายของคำศัพท์จาก

ศัพท์สัมพันธ์ของเวิร์ดเน็ต และการวิเคราะห์คำศัพท์ของเค้าร่างเมทาดาทา เพื่อกำหนดความสัมพันธ์ระหว่างเขตข้อมูลแต่ละเขตข้อมูล โดยกระบวนการนี้จะพิจารณาจากความสัมพันธ์ระหว่างคลาสหรืออ็อบเจกต์พร็อพเพอร์ตี้ (Object Properties) ที่เหมือนกันแต่ชื่อคลาสมีความแตกต่างกันอย่างสิ้นเชิง หรือมีค่าคะแนนความคล้ายคลึงต่ำกว่า 0.7 รวมถึงคลาสที่มีการจัดเก็บด้วยโครงสร้างของคลาสที่ซับซ้อน หรือคลาสที่มีโครงสร้างหลายระดับ ตามรูปที่ 5.8



รูปที่ 5.8 ลักษณะความขัดแย้งเชิงโครงสร้างสำหรับข้อมูลผู้แต่ง Writer และ Author

จากรูปที่ 5.8 แสดงให้เห็นถึงคลาสผู้แต่ง (Writer) ซึ่งบางครั้งสามารถใช้คำศัพท์ได้หลายคำ เช่น Author, Composer หรือเป็นกลุ่มผู้รับผิดชอบและเผยแพร่เนื้อหา (Contributor) และเมื่อวัดความคล้ายคลึงเชิงความหมายจากโครงสร้างศัพท์สัมพันธ์ของเวิร์ดเน็ตแล้วมีระยะทางที่ห่างกันมากทำให้คะแนนค่าความคล้ายคลึงต่ำ ก็จะใช้การพิจารณาอ็อบเจกต์พร็อพเพอร์ตี้ที่เหมือนกัน หรือสอดคล้องใกล้เคียงกันกัน และเพิ่มความสัมพันธ์ด้วยคุณสมบัติของภาษาอาว (OWL) อาทิ owl:equivalentClass owl:hasValue owl:someValuesFrom owl:allValuesFrom เข้ามาช่วยในการเชื่อมความสัมพันธ์ดังกล่าว ดังแสดงในรูปที่ 5.9



รูปที่ 5.9 ตัวอย่างการแก้ปัญหาความขัดแย้งเชิงโครงสร้าง

จากรูปที่ 5.9 สามารถนำคุณสมบัติ owl:allValueFrom ของภาษา OWL มาใช้ในการเชื่อมความสัมพันธ์ระหว่างคลาส writer เข้ากับคลาสอื่น ๆ ที่มีคุณสมบัติ createdBy เพื่อเป็นการจัดเก็บข้อมูลจากคลาส Contributor และ Author ร่วมกันได้ เนื่องจากทั้ง 2 คลาสนี้มีข้อมูลที่จัดเก็บคล้ายคลึงกันในลักษณะโครงสร้างการจัดเก็บข้อมูล

(4) การสร้างชุดข้อมูลอาร์ดีเอฟ

จากสร้างชุดข้อมูลอาร์ดีเอฟ จะใช้เครื่องมือในการช่วยผสานโครงสร้าง ชื่อว่า D2RQ platform (D2RQ, 2012) โดยผู้วิจัยทำการสร้างไฟล์กำหนดเงื่อนไขการผสานคลาสและเมทาดาทาตามการออกแบบดังกล่าวข้างต้น เพื่อให้การแปลงฐานข้อมูลและการผสานเมทาดาทามีความสะดวก รวดเร็ว

```
<vocab:contentobject_AstTypeID>AST0001</vocab:contentobject_AstTypeID>
<vocab:contentobject_Creator>เสถียร วิเชียรสาร</vocab:contentobject_Creator>
<vocab:contentobject_CreateDate>13ม.ค.2557</vocab:contentobject_CreateDate>
<vocab:contentobject_CoId>LO00038</vocab:contentobject_CoId>
<vocab:contentobject_LoTypeID>LOT0003</vocab:contentobject_LoTypeID>
<vocab:contentobject_CoTitle>ฟังก์ชั่นตรีโกณมิติ</vocab:contentobject_CoTitle>
<vocab:contentobject_ConDesc>สื่อประกอบการสอนเรื่องฟังก์ชั่นตรีโกณมิติ ประกอบด้วย วงกลมหนึ่งหน่วย
ฟังชั่นไซน์และโคไซน์ ฟังก์ชั่นตรีโกณมิติอื่นๆ ตัวผกผันของฟังก์ชั่นตรีโกณมิติ
กฎของโคไซน์และไซน์</vocab:contentobject_ConDesc>
<vocab:contentobject_Objective>นักเรียนมีความเข้าใจในฟังก์ชั่นตรีโกณมิติ
สามารถแก้โจทย์ปัญหาโดยใช้หลักการของตรีโกณมิติได้ รวมถึงการคำนวณฟังก์ชั่นไซน์
และโคไซน์ได้</vocab:contentobject_Objective>
```

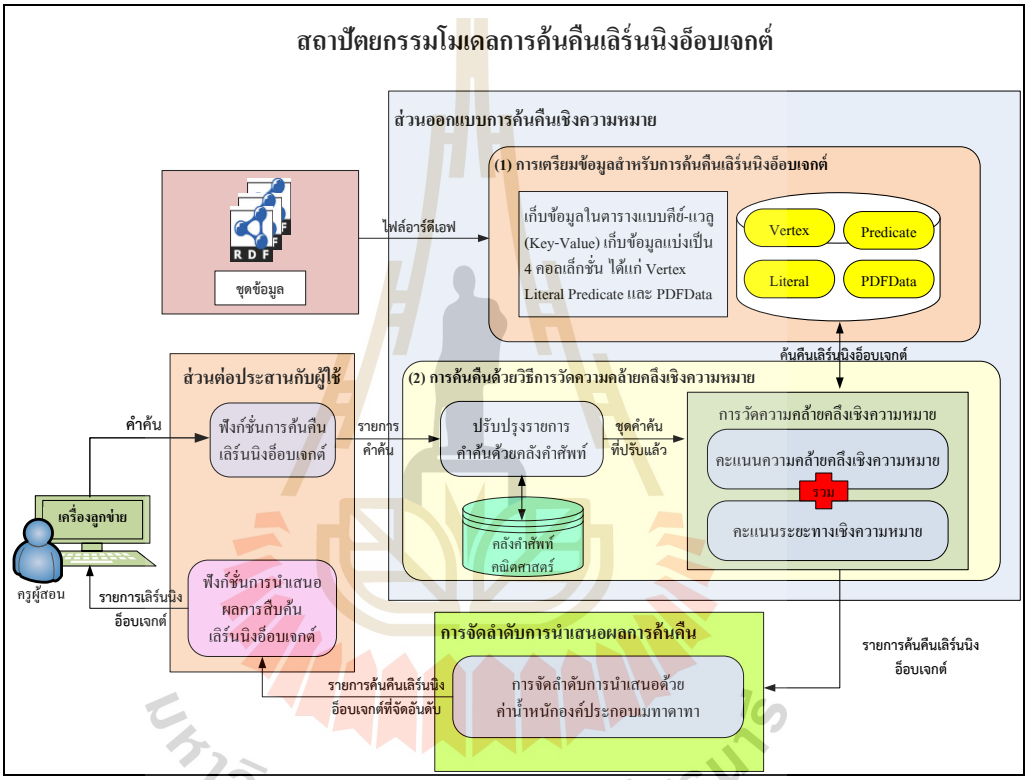
รูปที่ 5.10 ตัวอย่างไฟล์อาร์ดีเอฟหลังจากการผสานเมทาดาทาและแปลงโครงสร้าง

จากรูปที่ 5.10 เป็นชุดข้อมูลอาร์ดีเอฟที่แปลงจากฐานข้อมูลเชิงสัมพันธ์ ซึ่งในตัวอย่างเป็นคลาสข้อมูลย่อย (Asset) ที่เก็บไฟล์เลิร์นนิงอ็อบเจกต์รหัส LO0003 เรื่องฟังก์ชั่นตรีโกณมิติ ให้อยู่ในรูปแบบคลาสเลิร์นนิงอ็อบเจกต์ และกำหนดให้แอตทริบิวต์ในตารางข้อมูลให้เป็นดาทาพร็อพเพอร์ตี้ (data properties) ตามรูปแบบโครงสร้างของอาร์ดีเอฟ

5.2 ผลการออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์

โมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ มีองค์ประกอบส่วนงานทั้งหมด 2 ส่วน ได้แก่ (1) การออกแบบการค้นเชิงความหมาย เป็นขั้นตอนในการค้นคืนเชิงความหมาย โดยทำการเตรียมชุด

ข้อมูลอาร์ดีเอฟของเลิร์นนิงอ็อบเจกต์ที่จัดเก็บไว้ มาปรับให้เป็นรูปแบบตารางข้อมูลเพื่อความรวดเร็วในการเข้าถึงข้อมูล จากนั้นในขั้นตอนการค้นคืนจะใช้การวิเคราะห์เชิงความหมายประกอบด้วยการคิดคะแนนความคล้ายคลึงเชิงความหมายและการคิดคะแนนระยะทางเชิงความหมายนำมาประเมินขอบเขตของการค้นคืน (2) การจัดลำดับการนำเสนอผลการสืบค้น โดยใช้แนวคิดการให้น้ำหนักข้อมูลและคำศัพท์เมทาดาทาที่ผู้ใช้เลือกในการคิดคะแนนในการจัดลำดับผลการสืบค้น โดยแสดงสถาปัตยกรรมโมเดลดังรูปที่ 5.11



รูปที่ 5.11 สถาปัตยกรรมโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์

5.2.1 การออกแบบการค้นคืนเชิงความหมาย

ในขั้นตอนนี้จะนำเสนอกระบวนการและเทคนิคการสืบค้นโดยใช้แนวคิดการวัดความคล้ายคลึงเชิงความหมายเข้ามาช่วยในการเปรียบเทียบคำค้นในชุดข้อมูลอาร์ดีเอฟ ซึ่งกระบวนการสืบค้นเลิร์นนิงอ็อบเจกต์จะแบ่งออกเป็น 2 กระบวนการ ได้แก่ (1) การเตรียมข้อมูลสำหรับการค้นคืนเลิร์นนิงอ็อบเจกต์ เป็นการนำชุดข้อมูลอาร์ดีเอฟมาเตรียมข้อมูลในรูปแบบตารางข้อมูลและกำหนดโครงสร้างแบบคีย์-แวลู (Key-Value) เพื่อช่วยให้การเข้าถึงข้อมูลทำได้รวดเร็วแทนการเข้าถึงข้อมูลจากชุดข้อมูลอาร์ดีเอฟโดยตรงซึ่งจะใช้เวลาในการเข้าถึงข้อมูลทีละทริเปิล ทำให้ใช้

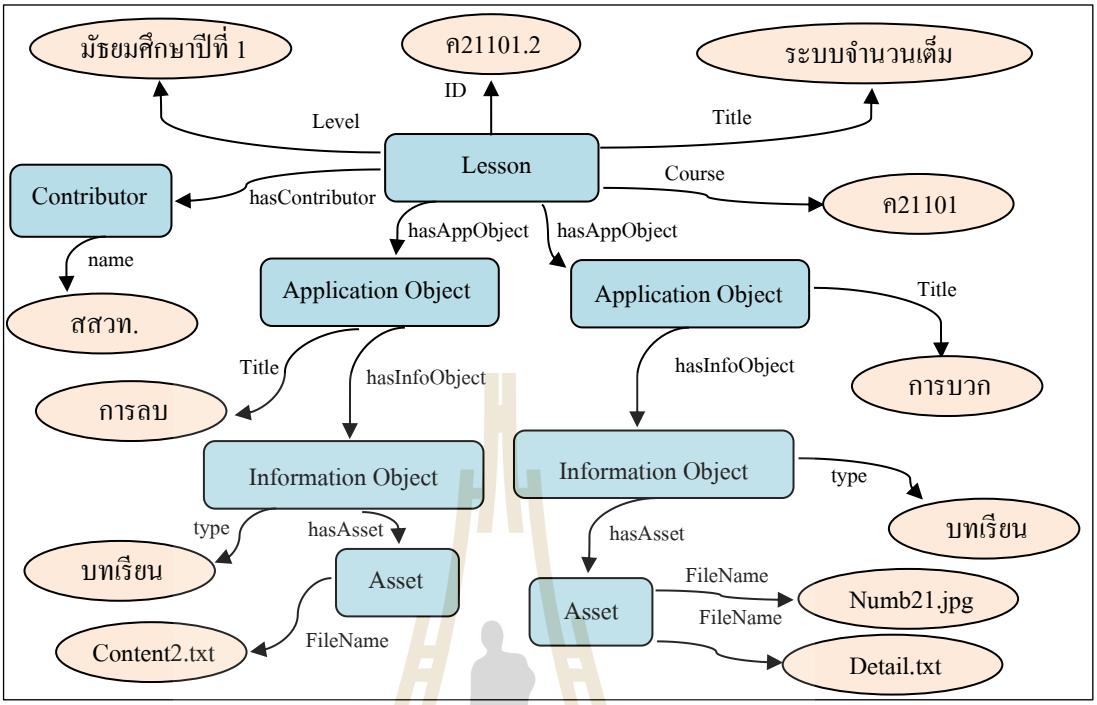
เวลาในการค้นคืนนานกว่าการจัดเก็บแบบตารางข้อมูล (2) การคำนวณค่าคะแนนความคล้ายคลึงเชิงความหมาย เป็นกระบวนการคัดเลือกรายการเลิร์นนิงอ็อบเจกต์ที่เกี่ยวข้องกับคำค้น โดยมีรายละเอียดดังต่อไปนี้

(1) การเตรียมข้อมูลสำหรับการค้นคืนเลิร์นนิงอ็อบเจกต์

เป็นกระบวนการกำหนดโครงสร้างดัชนีให้กับชุดข้อมูลเลิร์นนิงอ็อบเจกต์ที่ได้จากการแปลงจากระบบฐานข้อมูลเชิงสัมพันธ์ โดยกระบวนการเตรียมข้อมูล (Data preparation) ก่อนดำเนินการกำหนดโครงสร้างดัชนีให้กับชุดข้อมูล คือการกำหนดข้อมูลตัวแทน (ID) ให้กับข้อมูลแต่ละโหนด โดยใช้หลัก Key Values Design (Termehchy and Winslett, 2010) เพื่อให้สามารถขยายขอบเขตการทำโครงสร้างดัชนีเพิ่มเติมข้อมูลในอนาคตได้

โครงสร้างดัชนีแบ่งข้อมูลออกเป็น 4 คอลเล็กชัน (Bae, Nguyen, Kang, and Oh, 2013) โดยประยุกต์แนวการทำดัชนีแบบ Key Values Design ได้แก่ (1) คอลเล็กชั่นคลาสที่เป็นซับเจกต์ (Subject) และคลาสที่เป็นอ็อบเจกต์ (Object) โดยมีจุดยอด (Vertex) เพื่อเก็บข้อมูลซับเจกต์และอ็อบเจกต์ที่คู่กัน เพื่อแยกจุดยอดของโครงสร้างอาร์อีเอฟทั้งหมดไว้ที่เดียวกันเพื่อความรวดเร็วในการเข้าถึงโหนดเริ่มต้น (2) คอลเล็กชั่นชุดข้อมูลคุณสมบัติของทรัพยากร (Predicate) หรืออ็อบเจกต์พร็อพเพอร์ตี้ (Object Properties) (3) คอลเล็กชั่นดาทาพร็อพเพอร์ตี้ (Data Properties) หรือค่าของคุณสมบัติประเภทอักขระหรือลิตเทอร์รัล (Literal) สำหรับเป็นตารางในการเปรียบเทียบคำค้น 4) ข้อมูลอาร์ดีเอฟ (RDF Data Collection) เก็บดัชนีความสัมพันธ์ระหว่างชุดข้อมูลทั้ง 3 คอลเล็กชั่น เพื่อใช้ในการเข้าถึงคอลเล็กชั่นด้วยดัชนีและเชื่อมโยงรายละเอียดจากคอลเล็กชั่นทั้ง 3 เปรียบเสมือนตารางบอกเส้นทางที่จะเป็นจุดเริ่มต้นในการค้นคืนข้อมูลในชุดข้อมูลอาร์ดีเอฟทั้งหมด โดยขอยกตัวอย่างโครงสร้างอาร์ดีเอฟเลิร์นนิงอ็อบเจกต์เพื่ออธิบายกระบวนการจัดทำโครงสร้างดัชนีเอกสาร

มหาวิทยาลัยเทคโนโลยีสุรนารี



รูปที่ 5.12 ตัวอย่างกราฟอาร์ดีเอฟเลิร์นนิงอ็อบเจกต์หัวข้อระบบจำนวนเต็ม

จากภาพ 5.12 เป็นตัวอย่างโครงสร้างอาร์ดีเอฟเลิร์นนิงอ็อบเจกต์ในบทเรียนเรื่องระบบจำนวนเต็ม ซึ่งมีการกำหนดให้มีคลาสบทเรียน (Lesson) ประกอบด้วยดาทาพร็อพเพอร์ตี (Data Properties) รหัสบทเรียน (ID) ชื่อบท (Title) ระดับชั้นปี (Level) และรหัสรายวิชา (Course) โดยจะมีอ็อบเจกต์พร็อพเพอร์ตีกับคลาสผู้แต่ง (Contributor) คลาสเนื้อหาการเรียนรู้ (Application Object) ในคลาสผู้แต่ง (Contributor) มีดาทาพร็อพเพอร์ตีชื่อผู้แต่ง (name) และคลาสเนื้อหาการเรียนรู้ (Application Object) มีดาทาพร็อพเพอร์ตีชื่อเรื่อง (Title) และอ็อบเจกต์พร็อพเพอร์ตี hasInfoObject เพื่อเชื่อมโยงรายละเอียดเลิร์นนิงอ็อบเจกต์ในระดับวัตถุสารสนเทศ (Information Object) ซึ่งจะมีดาทาพร็อพเพอร์ตีประเภทสื่อ (Type) และอ็อบเจกต์พร็อพเพอร์ตี hasAsset เพื่อเชื่อมโยงกับไฟล์ข้อมูลในเลิร์นนิงอ็อบเจกต์นั้น ๆ จากคลาสข้อมูลย่อย (Asset) ซึ่งจะมีดาทาพร็อพเพอร์ตีชื่อไฟล์ (FileName) เพื่อจำแนกไฟล์องค์ประกอบย่อยภายในเลิร์นนิงอ็อบเจกต์ให้สามารถนำไปใช้งานได้ จากโครงสร้างดังกล่าว สามารถนำมาเตรียมข้อมูลเพื่อปรับเป็นโครงสร้างดัชนีชุดข้อมูลได้ โดยการแบ่งเป็นคอลเล็กชั่นของดัชนี ดังต่อไปนี้

คอลเล็กชั่น (Collection) เป็นการปรับรูปแบบของโครงสร้างอาร์ดีเอฟในหนึ่งชุดข้อมูลให้เก็บในรูปแบบตารางตามแนวคิดการเตรียมข้อมูลก่อนการประมวลผล (Pre Processing)

โดยการเก็บข้อมูลเป็น 4 คอลเล็กชั่นตามการอธิบายเบื้องต้น ซึ่งชุดข้อมูลจะถูกแปลงให้อยู่ในรูปแบบทริเปิล ตามตารางที่ 5.11

ตารางที่ 5.11 ตัวอย่างข้อมูลทริเปิลของเลิร์นนิงอ็อบเจกต์

Subject	Predicate	Object
Lesson	ID	ค21101.2
Lesson	Title	ระบบจำนวนเต็ม
Lesson	Level	มัธยมศึกษาปีที่ 1
Lesson	Course	ค21101
Lesson	hasContributor	Contributor
Lesson	hasAppObject	Application Object
Contributor	name	สสวท.
Application Object	hasInfoObject	Information Object
Application Object	Title	การบวก
Application Object	Title	การลบ
Information Object	Type	บทเรียน
Information Object	Type	บทเรียน
Information Object	hasAsset	Asset
Asset	FileName	Deatil.txt
Asset	Filename	Context2.txt
Asset	Filename	Numb21.jpg

จากตารางที่ 5.11 ข้อมูลตัวอย่างทริเปิลสามารถวิเคราะห์ระยะทางระหว่างคำศัพท์ของแต่ละคลาสและข้อมูลค่าของคุณสมบัติประเภทอักขระหรือเรียกว่าลิตเทอร์รัล (Literal) ได้ โดยระยะห่างระหว่างโหนด Lesson กับโหนดข้างเคียงที่มีค่า K-Depth เท่ากับ 1 คือ “ค21101.1” “ระบบจำนวนเต็ม” “มัธยมศึกษาปีที่ 1” “Contributor” “Application Object” และ “ค21101” ดังนั้นจากแนวคิดการกำหนดดัชนีด้วย Key: Values Design ผู้วิจัยจะออกแบบการแทนค่าข้อมูลในตารางด้วยตัวเลขดัชนี (Vertex) ของจุดยอดในแต่ละคู่ของข้อมูล โดยสร้างเป็นคอลเล็กชั่นที่ 2 ตารางคอลเล็กชั่นจุดยอด (Vertex Collection) เพื่อเก็บข้อมูลคู่ของซับเจ็กต์และอ็อบเจกต์ในรูปแบบ

(vertex ID, <incoming node, outgoing node>) (5-2)

จากรูปแบบของการจัดเก็บข้อมูลที่ 5.12 คู่ข้อมูลดังกล่าวจะมีรหัสจุดยอด (vertex ID) เพื่อเป็นจุดอ้างอิงคู่ลำดับข้อมูลแต่ละชุด ในคู่ลำดับจะเก็บข้อมูลของโหนดที่ส่งเข้า และโหนดที่ส่งออก โดยในโหนดนำเข้า (incoming node) จะมีการระบุรหัสคุณสมบัติของทรัพยากร (Predicate) เพื่อบอกรูปแบบความสัมพันธ์ของคู่โหนดนั้น โดยใช้เลขรหัสจากคอลเล็กชั่นจุดยอดที่มีรายละเอียดตามตารางที่ 5.12 และคอลเล็กชั่นคุณสมบัติของทรัพยากร (Predicate Collection) สามารถระบุได้ตามตารางที่ 5.13 และได้ตารางลำดับคู่โหนดในตารางที่ 5.14 ตามลำดับ

ตารางที่ 5.12 คอลเล็กชันจุดยอด (Vertex Collection)

VID	Value
0	Lesson
1	Contributor
2	ค21101
3	Application Object
4	Information Object
5	Asset
6	ค21101.2
7	ระบบจำนวนเต็ม
8	ชั้นมัธยมศึกษาปีที่ 1
9	สสวท.
10	บทเรียน
11	การบวก
12	การลบ
13	Content2.txt
14	Detail.txt
15	Numb21.jpg

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตารางที่ 5.13 คอลเล็กชันคุณสมบัติของทรัพยากร (Predicate Collection)

PID	Value
0	ID
1	Level
2	Title
3	Course
4	hasContributor
5	hasAppObject
6	hasInfoObject
7	hasAsset
8	type
9	Name
10	FileName

ตารางที่ 5.14 คอลเล็กชันอาร์ดีเอฟเลิร์นนิงอ็อบเจกต์ (RDF Collection of Learning Object)

VID	Incoming Node	Outgoing Node
[0]	{}	[1,2,3,5,6,7]
[1]	{“4”: [0]}	[9]
[2]	{“3”: [0]}	[]
[3]	{“5”: [0]}	[4,10]
[4]	{“6”: [3]}	[11]
[5]	{“0”: [0]}	[]
[6]	{“2”: [0]}	[]
[7]	{“8”: [1]}	[]
[8]	{“0”: [2]}	[]
[11]	{“3”: [2]}	[]

จากตารางที่ 5.12 เป็นการแสดงการจับคู่โหนดนำเข้าและส่งออกตามโครงสร้างอาร์ดีเอฟที่ออกแบบไว้จากรูปที่ 5.12 เพื่อกำหนดโครงสร้างดัชนีให้กับโหนดแต่ละโหนด เช่น VID

ที่ 0 คือโหนดบทเรียน (Lesson) จะไม่มีการนำเข้าข้อมูลจากโหนดอื่น ๆ จึงเป็นค่าว่าง แต่มีความสัมพันธ์กับโหนดอื่น ๆ อีก 6 โหนดคือโหนดที่ 1 ผู้แต่ง (Contributor) โหนดที่ 2 รายวิชา ค 21101 โหนดที่ 3 เนื้อหาการเรียนรู้ (Application Object) โหนดที่ 4 “ค21101.1” โหนดที่ 5 “ระบบจำนวนเต็ม” และโหนดที่ 6 “มัธยมศึกษาปีที่ 1” หรือที่โหนดที่ 3 คลาส “Application Object” จะมีโหนดนำเข้าคือ โหนด 0 บทเรียน (Lesson) ด้วยคุณสมบัติ hasAppObject (PID=5) จึงเขียนเป็น {5,[0]} และมีความสัมพันธ์ส่งออกไป 2 โหนด คือโหนดที่ 4 วัตถุสารสนเทศ (Information Object) VID=4 และโหนดที่ 2 “บทเรียน” VID=10 เป็นต้น

เมื่อทำการเตรียมข้อมูลของเลิร์นนิงอ็อบเจกต์ในรูปแบบตารางที่กำหนดแล้ว จะทำการสืบค้นข้อมูลข้างต้นได้ เมื่อผู้ใช้ใส่คำค้นเข้าสู่ระบบและทำการปรับปรุงคำค้นโดยคลังคำทางคณิตศาสตร์ ระบบจะทำการเปรียบเทียบคำค้นกับคลังคำศัพท์คณิตศาสตร์ที่สร้างขึ้น (บทที่ 3 หน้า 107) เพื่อหาคำค้นที่เหมาะสมและหัวเรื่องที่เกี่ยวข้อง โดยขอยกตัวอย่างชุดคำศัพท์ที่เกี่ยวข้องกับตัวอย่างข้างต้น ตามภาพที่ 5.13

คณิตศาสตร์				
คณิตศาสตร์	จำนวนและการคำนวณ			
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน		
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน	จำนวนนับ	
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน	จำนวนเต็ม	
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน	จำนวนตรรกยะ	
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน	จำนวนอตรรกยะ	
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน	จำนวนพิชคณิต	
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน	จำนวนจริง	
คณิตศาสตร์	จำนวนและการคำนวณ	แนวคิดจำนวน	จำนวนเชิงซ้อน	

คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต		
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	การบวก
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	การลบ
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	การคูณ
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	การหาร
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	ราก
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	แฟกทอเรียล
คณิตศาสตร์	จำนวนและการคำนวณ	เลขคณิต	การดำเนินการ	การประมาณ

ภาพที่ 5.13 โครงสร้างคำศัพท์ทางคณิตศาสตร์เรื่องจำนวนและการคำนวณ

จากภาพที่ 15.3 เมื่อผู้ใช้ใส่คำค้น “การบวก จำนวนเต็ม” ระบบจะทำการเปรียบเทียบกับคลังคำ ซึ่งจะตรงกับกลุ่มคำศัพท์ภาษาไทย ในหัวเรื่อง {คณิตศาสตร์, จำนวนและการคำนวณ, แนวคิดจำนวน, จำนวนเต็ม} และ {คณิตศาสตร์, จำนวนและการคำนวณ, เลขคณิต, การดำเนินการ, การบวก} ดังนั้นจากกลุ่มคำที่เกี่ยวข้องจะถูกดึงมาเป็นคำค้นเพื่อให้การค้นครอบคลุมเรื่องอื่น ๆ จากหัวเรื่องได้ โดยจะทำการเลือกคำค้นในแบบสองภาษา ดังนั้นคำค้นที่ได้จะมี 2 ชุด ได้แก่

{คณิตศาสตร์, จำนวนและการคำนวณ, เลขคณิต, แนวคิดจำนวน, จำนวนเต็ม, การบวก}

{mathematic, number and computation, arithmetic, number concept, integer, addition}

เมื่อได้รายการคำค้นจะทำการสืบค้นโดยเปรียบเทียบคำค้น กับตารางคอลเล็กชั่นของโหนดที่เตรียมไว้ โดยการเปรียบเทียบกับตารางคอลเล็กชั่นจุดยอด (Vertex Collection) เพื่อหาข้อมูลที่มีความคล้ายคลึงกับคำค้นที่ใส่เข้าไป จะได้โหนดที่มีความคล้ายคลึงหรือตรงกับคำค้น ดังตารางที่ 5.15

ตารางที่ 5.15 คอลเล็กชันจุดยอดที่เกี่ยวข้องกับคำค้น

VID	Value
7	ระบบจำนวนเต็ม
11	การบวก
0	Lesson
3	Application Object

จากคอลเล็กชั่นจุดยอดจะสามารถเชื่อมโยงไปยังคุณสมบัติของทรัพยากร ในคลาสที่เกี่ยวข้องได้ ตามตารางที่ 5.16

ตารางที่ 5.16 คอลเล็กชันคุณสมบัติของทรัพยากรที่เกี่ยวข้องกับคำค้น

PID	Value
2	Title

จากตารางที่ 5.15 และ 5.16 จะทำให้ทราบว่าระบบจะทำการค้นคืนโดยเริ่มต้นที่คลาส Lesson และ Application Object ที่มีข้อมูลของคุณสมบัติทรัพยากร “Title” เป็น “ระบบจำนวนเต็ม และการบวก” โดยคลาสเหล่านั้นจะเป็นจุดเริ่มเพื่อทำการเปรียบเทียบคำค้นและรายละเอียดภายในคลาส รวมถึงการเชื่อมโยงไปยังคลาสอื่น ผ่านอ็อบเจกต์พร็อพเพอร์ตีที่เชื่อมกับคลาสข้างต้น ได้แก่ hasAppObject และ hasInfoObject เพื่อเปรียบเทียบข้อมูลในคลาสที่เกี่ยวข้องต่อไป ทั้งนี้การเปรียบเทียบจะใช้การคำควณค่าคะแนนความคล้ายคลึงเชิงความหมาย เพื่อประเมินว่าคลาสใดบ้างที่มีข้อมูลที่มีความคล้ายคลึงกับคำค้นมากที่สุด

(2) การคำนวณค่าคะแนนความคล้ายคลึงเชิงความหมาย (Semantic Similarity Score: SSS)

จากปัญหาของการสืบค้นเลิร์นนิงอ็อบเจกต์ในบทที่ 2 หัวข้อ 2.3.2 กล่าวคือ การสืบค้นด้วยคำค้นทั่วไปในโอเพนดาทาจะมีปัญหาเรื่องความแม่นยำต่ำ เนื่องจากลักษณะข้อมูลที่เก็บในชุดข้อมูลสามารถเชื่อมโยงไปยังชุดข้อมูลอื่น ๆ ได้ ทำให้ขอบเขตการสืบค้นด้วยคำค้นเพียงคำเดียวสามารถดึงข้อมูลที่เกี่ยวข้องได้ในจำนวนที่มากเกินจริง ดังนั้นจึงมีงานวิจัยหลายชิ้นที่คิดค้นวิธีกระบวนการสืบค้นโดยการกำหนดขอบเขตหรือระดับความลึก (K-Depth) ในการสืบค้นข้อมูล ซึ่งจะเกิดปัญหาของการกำหนดระดับความลึกในการสืบค้นที่ไม่เหมาะสม ได้แก่การกำหนดขอบเขตการค้นที่แคบเกินไป การเปรียบเทียบคำค้นมีความแม่นยำมาก แต่ผลลัพธ์ที่เกี่ยวข้องจะไม่ถูกสืบค้นขึ้นมาได้ครบถ้วน หรือการกำหนดความลึกมากเกินไปก็จะส่งผลต่อจำนวนผลลัพธ์ที่ออกมามากเกินความจำเป็น และส่งผลต่อประสิทธิภาพของระบบอย่างมาก เป็นต้น ดังนั้นในกระบวนการนี้จะเป็นการใช้การวัดความคล้ายคลึงเชิงความหมายช่วยในการกำหนดระดับความลึกในการสืบค้นภายในโครงสร้างอาร์ดีเอฟที่เหมาะสมที่สุด

กระบวนการวัดความคล้ายคลึงเชิงความหมายจะทำการคำนวณค่าคะแนนความคล้ายคลึงจาก 2 ส่วน ได้แก่

(2.1) การวิเคราะห์ความสัมพันธ์ของคุณสมบัติของทรัพยากร (Relational Predicate Keyword) จากรูปที่ 5.13 คลาสแต่ละคลาสจะสัมพันธ์กันด้วยคุณสมบัติของทรัพยากร โดยจะมีคำศัพท์ที่ใช้บ่งชี้ความสัมพันธ์ เช่น ID, Title, Level เป็นต้น ซึ่งคำศัพท์เหล่านี้สามารถบ่งบอกถึงลักษณะความสัมพันธ์ของตัวทรัพยากร (Subject) และค่าของคุณสมบัติ (Object) แต่ละตัวได้ และบ่งบอกถึงขอบเขตและเนื้อหาของกลุ่มคำศัพท์ในเชิงความหมายได้ การวัดความคล้ายคลึงนี้จะได้แนวคิดการวัดความคล้ายคลึงของลิน (Lin, 1998) เนื่องจากเป็นการวัดความคล้ายคลึงของข้อมูลในโครงสร้างที่มีความสัมพันธ์ระหว่างคำศัพท์

$$sim_{Lin} = \frac{2.IC(LCS)}{IC(concept_1)+IC(concept_2)} \qquad (5\text{-}3)$$

โดย

IC คือ ค่า Information Concept โดยคิดจากความน่าจะเป็นที่ปรากฏ concept ทั้งสอง

LCS คือ คำที่เชื่อมกันใกล้สุดระหว่าง concept ที่ 1 แล2

กระบวนการในการวัดความคล้ายคลึง เริ่มจากผู้ใช้ใส่คำสำคัญในการค้นเข้ามา จะนำคำค้นนั้นไปเปรียบเทียบกับโหนดที่เป็นค่าของคุณสมบัติของทรัพยากรแบบอักขระ (Literal) การเปรียบเทียบใช้การเปรียบเทียบแบบจับคู่ (Matching) เพื่อให้ได้ค่าที่ตรงกับคำค้นมากที่สุด และเก็บค่า Vertex ID ของค่าของคุณสมบัติประเภทอักขระ (Literal) นั้นเพื่อเชื่อมโยงกลับไปที่โหนดหลักผ่าน PID ที่เป็นรหัสของดาทาพร็อพเพอร์ตี้ (Data Properties) เมื่อเชื่อมข้อมูลสำเร็จก็จะได้ชุดของคำศัพท์คุณสมบัติ (List of Predicate) ที่จะนำไปใช้ในการวัดระยะทางของคำศัพท์ได้

เมื่อได้ชุดของคำศัพท์แล้วจะนำไปคำนวณหาคะแนนระยะทาง (Distance Score) โดยกำหนดให้จุดโหนดที่เป็นจุดเริ่มต้นมีค่าเท่ากับ 0 และเดินทางไปยังโหนดแต่ละระดับมีค่าเท่ากับ 1 ในแต่ละชั้นของโครงสร้าง โดยใช้สมการการหาคะแนนระยะทาง (5-4) โดยประยุกต์จากสูตรของแบ (Bae et al., 2013)

$$SDS(k, o_i) = 1 - \frac{NSC(k,o_i)-1}{Max[NSC(k,o_1)NSC(k,o_n)]} \qquad (5\text{-}4)$$

เมื่อ

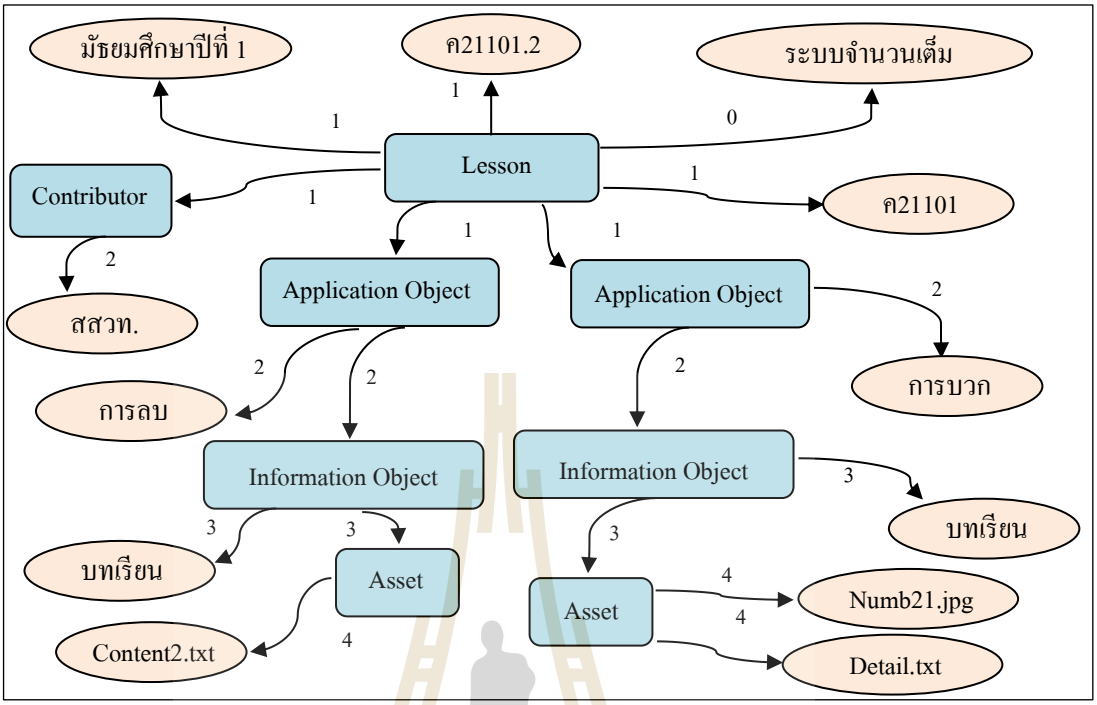
SDS คือ คะแนนระยะทางระหว่างคำค้นกับโหนด

k คือ คำค้นที่ใส่เข้าระบบ

o_i คือ ชุดข้อมูลที่เชื่อมโยงกันในอาร์ดีเอฟ

NSC คือ ค่าของการเปลี่ยนตำแหน่งของโหนดข้อมูล

จากสมการ 5-3 เป็นการหาค่าคะแนนระยะทาง โดยค่า Number of Subject Change: NSC จะพิจารณาจากจุดเริ่มต้นกำหนดให้คำค้นที่เปรียบเทียบแล้วตรงกับค่าของคุณสมบัติประเภทอักขระ (Literal) ตัวใด ให้นับโหนดนั้นมีค่าเป็น 0 จากนั้นจะทำการเคลื่อนย้ายตำแหน่งดัชนีไปตามคุณสมบัติของทรัพยากร (Predicate) แต่ละตัวที่เกี่ยวข้องกับโหนดหลัก จนถึงปลายโหนดค่าของคุณสมบัติประเภทอักขระ (Literal) อื่น ๆ จนครบ โดยนับจำนวนโหนดที่เดินทางผ่านเป็นค่า NSC จากรูปที่ 5.14 เป็นกระบวนการแทนค่าระยะทางในเส้นทางความสัมพันธ์ของโหนดแต่ละโหนด



รูปที่ 5.14 จำนวนคะแนนการเปลี่ยนโหนดโดยเริ่มที่รหัสบทเรียน “ระบบจำนวนเต็ม”

จากรูปที่ 5.14 เป็นการกำหนดคะแนนตามการเปลี่ยนโหนดของดัชนีในการท่องโหนดความสัมพันธ์ตามโครงสร้างและความสัมพันธ์ที่กำหนดไว้ ยกตัวอย่างการสืบค้นด้วยชื่อบทเรียน “ระบบจำนวนเต็ม” แล้ว ระบบจะท่องโหนดแต่ละโหนดไปตามขอบเขตจนถึงจุดปลายที่ค่าของคุณสมบัติประเภทอักขระ (Literal) ทุกโหนด ที่ค่าของคุณสมบัติ “บทเรียน” มีค่าระยะทางเท่ากับ 2 เนื่องจากต้องผ่านโหนดจำนวน 2 โหนดได้แก่ Lesson และ Application Object นั่นเอง

เมื่อทดสอบใช้สมการหาค่าคะแนนระยะทางเชิงความหมาย (Semantic Distance Score: SDS) ระหว่าง รหัสรายวิชา “ค21101.2” และ ชื่อบทเรียน “ระบบจำนวนเต็ม” โดยค่า NSC เท่ากับ 1 และค่า MAX[NSC] เท่ากับ 3 เนื่องจากพิจารณาทั้งโครงสร้าง ค่าของ NSC ที่มากที่สุดคือ 3 ที่โหนด “บทเรียน” ดังนั้น ค่า SDS จะมีคะแนนเท่ากับ 0.67 ตามรายละเอียดสมการ (5-4)

$$SDS(\text{“ค21101.2”}, \text{“ระบบจำนวนเต็ม”}) = 1 - ((1-1) / 3)$$

$$SDS(\text{“ค21101.2”}, \text{“ระบบจำนวนเต็ม”}) = 1$$

จากการวัดระยะทางเชิงความหมายจะมีค่าอยู่ระหว่าง 0 ถึง 1 โดย 0 หมายถึงไม่มีความเกี่ยวข้องหรือสอดคล้องกันเชิงความหมายเลย และ 1 หมายถึงมีความเกี่ยวข้องเชิงความหมายมากที่สุด หรือหมายถึงเป็นสิ่งเดียวกัน ใช้แทนกันได้ ดังนั้นจากตัวอย่างการคำนวณแสดงถึง รหัสบทเรียน “ค21101.2” มีความหมายเทียบเท่ากับชื่อเรื่อง “ระบบจำนวนเต็ม” เพราะเป็นสิ่งเดียวกันในเชิงความสัมพันธ์ ถึงแม้จะมีอักขระแตกต่างกันโดยสิ้นเชิงก็ตาม

(2.2) คะแนนการเปรียบเทียบเชิงความหมายระหว่างคลาส (Semantic Similarity Score) นอกจากการวัดระยะทางเชิงความหมายจากกระบวนการที่ 1 แล้ว ในงานวิจัยนี้ยังได้ใช้การเปรียบเทียบความคล้ายคลึงเชิงความหมายเพื่อนำไปคำนวณร่วมกับการวัดระยะทางเชิงความหมายเพื่อใช้ในการประเมินความเหมาะสมของระดับการสืบค้นทั้งรายละเอียดของคลาสกับคำค้น และระยะความลึกของการสืบค้นในโครงสร้างการจัดเก็บข้อมูลอีกด้วย ดังนั้นในส่วนการเปรียบเทียบเชิงความหมายของคำค้นกับคลาสแต่ละคลาสจะใช้การประยุกต์จากสูตรของลิน (Lin, 1998) ตามสมการที่ 5-5 และการเปลี่ยนตัวแปรจากการเปรียบเทียบคำค้นต่อคำในเอกสารให้เป็นคำค้นและอัตราส่วนของเขตข้อมูลในคลาสแต่ละคลาส ตามแนวคิดของแบ (Bae, Nguyen, Kang and Oh, 2013)

$$SSS(k, o_i) = \frac{2 \times \log r(c_o)}{\log r(c_i) + \log r(c_k)} \qquad (5\text{-}5)$$

เมื่อ

k คือ คำค้น

c_k คือ จำนวนของคลาสที่มีคำค้นปรากฏ

c_i คือ จำนวนคลาสที่มีอ็อบเจกต์ปรากฏ

c_o คือ จำนวนคลาสที่มีปรากฏคำค้นและอ็อบเจกต์

จากสมการที่ 5-4 เป็นการเปรียบเทียบความคล้ายของคำค้นกับคลาสที่มีค่าของคุณสมบัติประเภทอักขระ (Literal) ปรากฏอยู่ โดย $r(c_k)$ คืออัตราส่วนระหว่างคลาสที่มีคำค้นปรากฏอยู่กับจำนวนคลาสทั้งหมด $r(c_i)$ คืออัตราส่วนระหว่างคลาสที่มีค่าของคุณสมบัติของทรัพยากรที่ต้องการเปรียบเทียบปรากฏอยู่ กับจำนวนคลาสทั้งหมด และ $r(c_o)$ เป็นอัตราส่วนระหว่างจำนวนคลาสที่ปรากฏทั้งคำค้นและค่าของคุณสมบัติของทรัพยากรในคลาสนั้น กับจำนวนคลาสทั้งหมด โดยการเปรียบเทียบความคล้ายคลึงเชิงความหมายนี้จะทำการเปรียบเทียบจากคำค้นและโครงสร้างของคลังคำศัพท์ที่สร้างไว้ในขั้นตอนการสร้างคลังศัพท์สัมพันธ์

จากการหาระยะทางเชิงความหมายระหว่างคำค้น “ค21101.2” และโหนด “ระบบจำนวนเต็ม” สามารถนำมาเปรียบเทียบความคล้ายคลึงเชิงความหมายตามสมการที่ 5-4 ได้ โดยคลาสทั้งหมดในตัวอย่าง 5 คลาสได้แก่ Lesson, Contributor, Course, Learning Object และ Asset คลาสที่ปรากฏคำค้นมี 1 คลาส คือ Lesson คลาสที่ปรากฏโหนดที่จะเปรียบเทียบมี 1 คลาส คือ Lesson เมื่อคำนวณตามสมการแล้วจะได้ค่าเท่ากับ 1 คือทั้งสองคำมีความคล้ายคลึงเชิงความหมาย

SSS(“ค201101.2”, “ระบบจำนวนเต็ม”) = (2*log(1/5))/(log(1/5)+(log(1/5))

= 1

จากการวัดความคล้ายคลึงทั้งทางด้านระยะทางเชิงความหมาย (Semantic Distance Score: SDS) และการวัดความคล้ายคลึงเชิงความหมายระหว่างคลาส (Semantic Similarity Score: SSS) ผู้วิจัยจึงนำการวัดความคล้ายคลึงทั้งสองส่วนมาใช้ร่วมกัน โดยมีค่าพารามิเตอร์ (p) เพื่อเป็นสัมประสิทธิ์ในการกำหนดค่าผกผันระหว่างการวัดระยะทางเชิงความหมายและการวัดความคล้ายคลึงเชิงความหมายให้มีค่าอยู่ระหว่าง 0 ถึง 1 โดยการกำหนดค่าพารามิเตอร์ p ที่เหมาะสมเพื่อให้ประสิทธิภาพการสืบค้นมีค่าสูงที่สุด ซึ่งจะแสดงกระบวนการประเมินค่า p ที่เหมาะสมกับการสืบค้นในหัวข้อการประเมินผลประสิทธิภาพของโมเดลในลำดับถัดไป

$$SS(k, o_i) = p \cdot SDS(k, o_i) + (1 - p) \cdot SSS(k, o_i) \quad (5\text{-}6)$$

เมื่อ

k คือ คำค้น

o_i คือ อ็อบเจกต์ที่ทำการเปรียบเทียบ

p คือ ค่าพารามิเตอร์ให้น้ำหนัก

5.2.2 การจัดลำดับการนำเสนอผลการสืบค้น

หลังการคำนวณความคล้ายคลึงเชิงความหมายโดยมีส่วนประกอบทางด้านระยะทางเชิงความหมายและความคล้ายคลึงของคลาสเชิงความหมายแล้ว งานวิจัยนี้ยังได้เพิ่มการจัดการรายการผลการสืบค้นในส่วนการนำเสนอให้มีความสอดคล้องกับความต้องการของผู้ใช้ โดยวิเคราะห์เค้าร่างเมทาดาทาและข้อมูลตัวแทนเอกสารทางด้านสถิติของระเบียน เพื่อนำมาใช้ในการนำเสนอผลการสืบค้นได้ดียิ่งขึ้น

การจัดลำดับการนำเสนอในงานวิจัยนี้ประกอบด้วยแนวทางการจัดลำดับผลการสืบค้นโดยวิธีการ Query Dependent Ranking (Liu, 2009) ซึ่งเป็นการจัดลำดับผลการสืบค้นโดยยึดการเปรียบเทียบระหว่างคำค้นกับคำดัชนีตัวแทนของเอกสารเป็นสำคัญ พร้อมกันนี้ได้เพิ่มแนวคิดการเพิ่มค่าน้ำหนักของคำค้นและองค์ประกอบทางบรรณานุกรม (Bibliographic) เข้าไปเพื่อเปรียบเทียบผลการจัดลำดับกับการจัดลำดับโดยใช้ข้อมูลตัวแทนเอกสารแบบไม่มีการใช้ค่าน้ำหนัก และนำสมการคำนวณหาความคล้ายคลึงระหว่างข้อมูลสองรายการแบบปริภูมิเวกเตอร์ (Vector Space) ด้วยวิธีการวัดความคล้ายคลึงแบบโคซายน์ (Cosine Similarity) และการปรับปรุงให้รองรับการเพิ่มค่าน้ำหนักของปัจจัยบรรณานุกรมเพิ่มเติม (Zhou, Lawless, Min and Wade, 2010) ได้แก่

$$sim(d,q) = \frac{[\sum_{k=1}^{n} d*q]}{\sqrt{\sum_{k=1}^{n} d^2} * \sqrt{\sum_{k=1}^{n} q^2}} * F \qquad (5\text{-}7)$$

เมื่อ

d คือ ค่าน้ำหนักของโหนดที่พบ

q คือ ค่าน้ำหนักของคำสำคัญที่ใช้ค้นหา

F คือ ค่าพารามิเตอร์ให้น้ำหนักเมทาดาทา

จากสมการที่ 5-7 เป็นสมการหาความคล้ายคลึงของคำสองคำ โดยเปรียบเทียบความคล้ายเชิงเส้นคือไม่ยึดความหมายของคำแต่วัดความคล้ายคลึง โดยการให้ค่าน้ำหนักของโหนดที่พบในโครงสร้างอาร์ดีเอฟ และการให้น้ำหนักคำสำคัญที่ใช้สืบค้น นอกจากนี้ยังมีการให้ค่าน้ำหนักเมทาดาทาที่ต้องการสำหรับเรียงลำดับโดยการคูณด้วยค่าคงที่ ระหว่าง 0.1 ถึง 1.0 โดยเลือกให้น้ำหนักกับเมทาดาทาที่ต้องการ โดยผลการสำรวจความต้องการของผู้ใช้ พบว่าเขตข้อมูลที่ผู้ใช้นิยมใช้เพื่อพิจารณาเลือกผลการสืบค้นหรือจัดกลุ่มเนื้อหาการสืบค้น ได้แก่ (1) ชนิดของไฟล์เอกสาร เช่น เอกสารเวิร์ด รูปภาพ ไฟล์เสียง (2) ภาษาที่ใช้ในตัวสื่อ (3) ระดับชั้นปีการศึกษา (4) รูปแบบเอกสาร เช่น เอกสารฉบับเต็ม มีไฟล์แนบหรือไม่ (5) หัวเรื่องคณิตศาสตร์ที่อยู่ในเนื้อหา เป็นต้น ซึ่งข้อมูลเหล่านี้ช่วยให้ผู้ใช้ตัดสินใจได้ว่าจะเข้าไปดูรายละเอียดรายการสืบค้นนั้นหรือไม่ เพื่อลดเวลาการพิจารณา ผู้วิจัยจึงนำปัจจัยเหล่านี้มาเป็นปัจจัยที่ใช้ในการจัดลำดับการแสดงผลเพื่อให้ผู้ใช้พบรายการที่ตนเองต้องการรวดเร็วมากยิ่งขึ้น โดยการทดลองจะใช้ตัวอย่างผลการสืบค้นเลิร์นนิงอ็อบเจกต์ในตารางที่ 5.17 ในการประเมินการจัดลำดับผลการสืบค้น โดยตัวอย่างเป็นการใส่คำค้นประสมระหว่างคำค้นภาษาไทยและภาษาอังกฤษ สืบค้นข้อมูลเลิร์นนิงอ็อบเจกต์ โดยกำหนดให้ภาษาในเอกสารเป็นภาษาไทย 2 รายการและภาษาอังกฤษ 2 รายการ รวมจำนวน 4 รายการ ดังนั้นค่าปัจจัย (F) ที่ใช้พิจารณาคือภาษาที่จะมีผลต่อคะแนนความคล้ายคลึงของเอกสารดังต่อไปนี้

```
REPEAT
READ next คำกำหนดเงื่อนไขการจัดลำดับ (i)
CASE  LessonFormat (รูปแบบของบทเรียน)
Result1(i)= Search กลุ่มคำ (i) จากฐานข้อมูล
CASE ApplicationFormat (รูปแบบของเนื้อหาการเรียนรู้)
Result2(i)= Search กลุ่มคำ (i) จากฐานข้อมูล
CASE InformationFormat (รูปแบบของวัตถุสารสนเทศ)
Result3(i)= Search กลุ่ม (i) จากฐานข้อมูล
IF Result1(i) AND Result1(i+1) AND…. Result2(i) AND
Result2(i+1)
Result3(i) AND Result3(i+1)..not NULL
THEN
BestFormat = Result1(i) _ Result1(i+1) _ ….
Result2(i) _ Result2(i+1) _ ….
Result3(i) _ Result3(i+1) _ ….
ELSE
NormalFormat = Result2(i) + Result2(i+1) + … or
Result3(i) + Result3(i+1) + ….
WorseFormat = Result1(i) + Result1(i+1) + ….
END IF
END REPEAT
```

รูปที่ 5.15 การคิดค่าคะแนนเพื่อกำหนดลำดับการนำเสนอให้กับผลการสืบค้น

จากรูปที่ 5.15 เป็นอัลกอริทึมในการพิจารณาองค์ประกอบของผลการสืบค้นที่ผู้ใช้เลือกในการจัดลำดับเนื้อหา โดยจะให้คะแนนกับองค์ประกอบที่ผู้ใช้เลือกเพื่อให้รายการสืบค้นนั้นแสดงในลำดับต้น เมื่อผลการสืบค้นแต่ละรายการผ่านอัลกอริทึมนี้จะมีคะแนนดังแสดงตามตัวอย่างในตารางที่ 5.17

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตารางที่ 5.17 ตัวอย่างการวัดค่าความคล้ายคลึงที่ได้จากการสืบค้นเลิร์นนิงอ็อบเจกต์

รายการคำค้น\เอกสาร	LO1T	LO2E	LO3T	LO4E	F
เซต	1	1	0.3	0.3	1
Element	0.8	0.7	0.7	0.7	0.5
สับเซต	0.9	0.7	0.5	0.3	1
จำนวนจริง	0.5	0.5	0.8	0.8	1
Real numbers	0.5	0.5	0.8	0.8	0.5
สมการเชิงเส้น	0.4	0.5	1	0.8	1
Union	0.7	0.6	0.3	0.2	0.5
Exponential	0.2	0.3	1	1	0.5
สมการ	0.8	0.8	1	1	1

จากตารางที่ 5.17 เป็นตัวอย่างการวัดความคล้ายคลึง (Similarity Score) ที่จากการสืบค้นเลิร์นนิงอ็อบเจกต์ตัวอย่าง ผลการสืบค้นจะถูกนำมาจัดเรียงลำดับ โดยใช้การวัดความคล้ายคลึงด้วยโคซายน์จากสมการที่ 5-6 เพื่อวัดความคล้ายคลึงของคำค้น และโหนดที่มีความใกล้เคียงกับคำค้นนั้นปรากฏอยู่ โดยจะแบ่งเป็น 2 กลุ่มคือ กลุ่มที่ไม่มีการเพิ่มค่าน้ำหนักของภาษา (F) กับกลุ่มที่คิดค่าน้ำหนักของภาษาในเอกสาร เพื่อวัดความแตกต่างของคะแนนความคล้ายคลึง

ตารางที่ 5.18 เปรียบเทียบผลการวัดความคล้ายคลึงแบบมีการเพิ่มค่าน้ำหนักและไม่เพิ่มค่าน้ำหนัก

รายการเอกสาร	**ค่าความคล้ายคลึงไม่เพิ่มค่าน้ำหนัก**	**ค่าความคล้ายคลึงแบบเพิ่มค่าน้ำหนัก**
LO1T	0.27	0.90
LO2E	0.26	0.77
LO3T	0.27	0.89
LO4E	0.27	0.77

จากตารางที่ 5.18 แสดงผลการวัดความคล้ายคลึงของเอกสารกับคำค้น โดยกลุ่มผลลัพธ์แบ่งเป็น 2 กลุ่ม คือกลุ่มที่วัดความคล้ายคลึงโดยไม่มีการถ่วงน้ำหนักของคำสำคัญ และกลุ่มที่สองเป็นการวัดความคล้ายคลึงด้วยการถ่วงน้ำหนักของคำที่สอดคล้องกับภาษาของเลิร์นนิงอ็อบเจกต์ ซึ่งในกรณีตัวอย่างกำหนดให้เอกสารภาษาไทยมีความสำคัญมากกว่าเอกสาร

ภาษาอังกฤษ ดังนั้น ผลการทดลองแสดงให้เห็นว่ากลุ่มที่มีการเพิ่มค่าน้ำหนักให้คำสำคัญจะช่วยเพิ่มความแตกต่างให้ผลการคำนวณมากขึ้น โดย LO1T และ LO2T เป็นเอกสารในกลุ่มเนื้อหาเดียวกับ และ LO3T และ LO4T เป็นกลุ่มเอกสารเดียวกัน ถ้าพิจารณาการวัดความคล้ายคลึงโดยไม่ถ่วงค่าน้ำหนักแล้ว คะแนนเอกสารแต่ละรายการแทบไม่มีความแตกต่างกันมากนัก และแสดงถึงเอกสารภาษาไทยและภาษาอังกฤษมีค่าความสำคัญเท่าเทียมกัน แต่ถ้าใช้การเพิ่มค่าน้ำหนักให้คำสำคัญต่อเอกสารแล้ว การคำนวณจะแสดงถึงเอกสารภาษาไทยจะมีค่าความคล้ายคลึงมากกว่าเอกสารที่เป็นภาษาอังกฤษ ทั้งนี้ขึ้นอยู่กับค่าน้ำหนักที่กำหนดในการคำนวณ โดยในตัวอย่างใช้ค่าน้ำหนักที่ 0.5 ดังนั้น ความแตกต่างระหว่างเอกสารอาจจะมีน้อยทำให้ค่าคะแนนไม่แตกต่างอย่างชัดเจน ดังนั้น ถ้าต้องการจัดกลุ่มเอกสารที่มีความแตกต่างกันมากขึ้น ให้ปรับค่าน้ำหนักคำสำคัญในเมทาดาทาหรือข้อมูลที่ต้องการให้มากขึ้น ซึ่งจะส่งผลต่อคะแนนในภาพรวมได้

จากแนวคิดการจัดลำดับการนำเสนอผลการสืบค้นหรือปรับปรุงผลการสืบค้นจากองค์ประกอบของเลิร์นนิงอ็อบเจกต์ นำมาสู่การการกำหนดคำสั่งในการปรับปรุงผลการสืบค้นทางหน้าเว็บไซต์ดังภาพที่ 5.16 แสดงรายละเอียดองค์ประกอบของเลิร์นนิงอ็อบเจกต์ที่สามารถเลือกเพื่อปรับปรุงผลการสืบค้นให้แสดงรายการผลการสืบค้นตามองค์ประกอบที่ผู้ใช้ต้องการได้

รูปที่ 5.16 ตัวเลือกเพื่อจัดลำดับการนำเสนอผลการสืบค้น

5.3 ผลการประเมินโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์

การประเมินโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา ได้ทำการประเมิน 2 วิธี ได้แก่ (1) การประเมินประสิทธิภาพของระบบ ด้วยการทดสอบจากการวัดค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าเอฟเมเชอร์ (F-Measure) และการเปรียบเทียบ

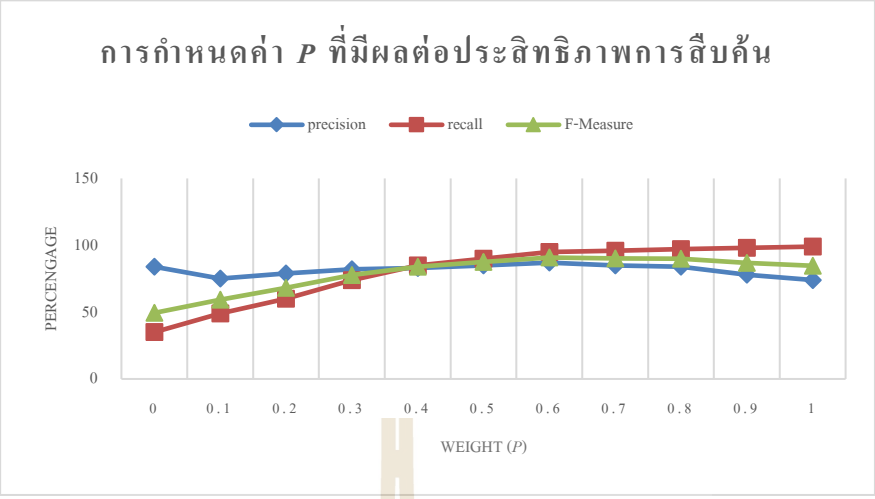
ความเร็วในการค้นคืนเลิร์นนิงอ็อบเจกต์กับเจนาเฟรมเวิร์ค (Jena Framework) และ (2) การประเมินการใช้งานระบบโดยผู้ใช้ ซึ่งมีรายละเอียดการประเมินผลดังนี้

5.3.1 ผลการประเมินประสิทธิภาพของระบบโดยการทดสอบการทำงานของและการเปรียบเทียบความเร็วในการค้นคืนเลิร์นนิงอ็อบเจกต์กับระบบอื่น

1) การประเมินประสิทธิภาพโดยการทำงานของระบบ

การทดสอบจะดำเนินการค้นคืนข้อมูลตัวอย่างที่ได้จากการแปลงข้อมูลจากฐานข้อมูลเชิงสัมพันธ์ ระเบียนข้อมูลจำนวน 300 รายการ และทำการเปรียบเทียบและแปลงเป็นโครงสร้างอาร์ดีเอฟ โดยการเตรียมข้อมูล (Pre-Processing) ให้อยู่ในรูปแบบตารางข้อมูลที่พร้อมทำการสืบค้นด้วยการจัดเก็บทริเปิลสโตร์ (Triple Store) จำนวน 18,000 ทริเปิล และพัฒนาเป็นระบบเว็บแอพลิเคชั่น โดยใช้ RAP API for RDF ในการพัฒนาเพื่อความสะดวกในการทดสอบระบบ โดยการทดสอบระบบจะนำเสนอเป็น 2 ส่วน ได้แก่ (1) การทดสอบกระบวนการเพื่อหาค่า p ที่เหมาะสมระหว่างการวัดความคล้ายคลึงเชิงความหมาย และการวัดระยะทางเชิงความหมาย (2) การเปรียบเทียบผลการสืบค้นระหว่างการสืบค้นโดยไม่ใช้ค่า p ในการสืบค้น และการสืบค้นที่ใช้ค่า p ที่เหมาะสมในการสืบค้น เพื่อนำไปทดสอบเพื่อหาค่าความแม่นยำ ค่าความระลึก และค่าเอฟเมเชอร์ โดยมีรายละเอียดดังนี้

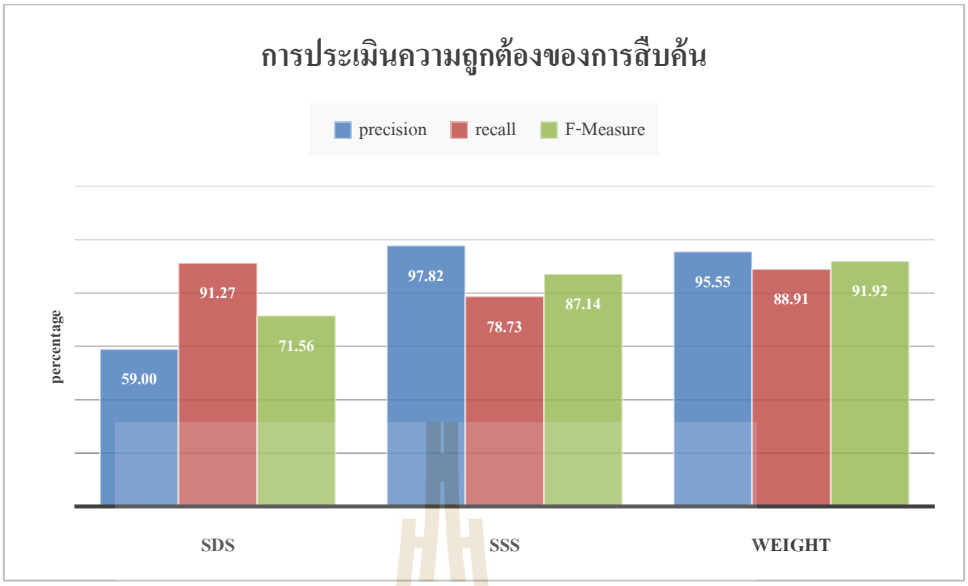
(1) การทดสอบกระบวนการเพื่อหาค่า P ในการวัดความคล้ายคลึง จากกระบวนการสืบค้นด้วยวิธีวัดความคล้ายคลึงเชิงความหมาย จากสมการที่ 5-5 จะมีค่าพารามิเตอร์ p สำหรับการคำนวณ เพื่อใช้เป็นค่าถ่วงน้ำหนักระหว่างค่าคะแนนระยะทางเชิงความหมาย (Semantic Distance Score: SDS) และค่าคะแนนความคล้ายคลึงเชิงความหมาย (Semantic Similarity Score: SSS) โดยค่า p ที่เหมาะสมนั้นจะเปลี่ยนแปลงไปตามโครงสร้างความสัมพันธ์ของชุดข้อมูลและลักษณะของทริเปิลในชุดข้อมูล ดังนั้นการกำหนดค่า p ที่เหมาะสมนั้นจะเป็นการปรับสมดุลของการวัดค่าความคล้ายคลึงและการวัดระยะทางเชิงความหมาย ดังนั้นจึงต้องมีการทดสอบการสืบค้นเพื่อหาค่า p ที่เหมาะสมของแต่ละชุดข้อมูล โดยผลการทดลองได้ค่า p แสดงดังรูปที่ 5.17



รูปที่ 5.17 กราฟแสดงค่าน้ำหนัก (p) ที่เหมาะสมกับชุดข้อมูล

จากรูปที่ 5.17 แสดงการทดสอบการสืบค้นโดยเปลี่ยนค่า p จาก 0 ถึง 1 โดยค่าที่เปลี่ยนแปลงไป จะมีผลต่อความแม่นยำ (Precision) และค่าความระลึก (Recall) โดยสองค่านี้เป็นค่าที่มีความสำคัญต่อการวัดประสิทธิภาพของการสืบค้นข้อมูลในโครงสร้างอาร์ดีเอฟ (RDF) อย่างมาก ซึ่งจะแสดงถึงผลการค้นที่มีความถูกต้องกับคำค้นที่ผู้ใช้ต้องการ และมีขอบเขตที่เหมาะสมได้เอกสารที่เกี่ยวข้องครบถ้วนมากที่สุด โดยในการทดลองนี้ค่า p ที่เหมาะสมและได้ค่าความแม่นยำและค่าความระลึกในระดับสูงคือ 0.4 ที่มีค่าความแม่นยำร้อยละ 87 ค่าความระลึกร้อยละ 95 และค่าเอฟเมเซอร์ (F-Measure) ที่ร้อยละ 90

(2) การเปรียบเทียบผลการสืบค้นโดยการสืบค้นโดยไม่ใช้ค่า p ทั้งนี้เมื่อนำข้อมูลค่าพารามิเตอร์ p มาใช้ในการสืบค้นเปรียบเทียบกับการสืบค้นด้วยการวัดระยะทางเชิงความหมายและการวัดความคล้ายคลึงเชิงความหมายเพียงอย่างเดียวโดยไม่นำค่า p มาใช้ในการคำนวณจะเห็นว่าประสิทธิภาพการใช้คะแนนแบบผสมจะให้ผลการสืบค้นที่มีประสิทธิภาพสูงกว่า ดังรูปที่ 5.18



รูปที่ 5.18 แผนภูมิสรุปการประเมินความถูกต้องของการสืบค้น

จากรูปที่ 5.18 แสดงถึงประสิทธิภาพการสืบค้นที่แตกต่างกันตามแต่ละวิธีโดยใช้ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าเอฟเมเชอร์ (F-Measure) จากการทดลองการสืบค้นด้วยคำค้นตั้งแต่ 1-5 คำและนำค่าเฉลี่ยมาแสดงผลในการประเมินประสิทธิภาพการสืบค้น ซึ่งแบ่งการทดสอบแบ่งเป็น 3 ลักษณะ ได้แก่

(1) วิธีการวัดระยะทางเชิงความหมาย (Semantic Distance Score: SDS) เป็นการสืบค้นโดยใช้คุณสมบัติของทรัพยากร (Predicate) เป็นคำศัพท์ในการพิจารณาถึงขอบเขตเรื่องที่สืบค้น โดยการวัดระยะทางระหว่างโหนดและคำค้นที่กำหนด ดังนั้น การใช้ระยะทางเชิงความหมายเพียงอย่างเดียว จะทำให้การสืบค้นสามารถสืบค้นผลลัพธ์ที่เกี่ยวข้องได้ค่อนข้างสูง ซึ่งพิจารณาจากค่าความถูกต้องหรือค่าการระลึก (Recall) สูงถึง 91.27 แสดงถึงผลการสืบค้นที่ค้นขึ้นมาได้นั้นมีความครบถ้วนมีเนื้อหาที่เกี่ยวข้องกับคำค้น เนื่องจากวิธีการนี้เป็นการสืบค้นตามความสัมพันธ์ ดังนั้น เมื่อชุดข้อมูลมีการเชื่อมความสัมพันธ์ในกลุ่มเดียวกันหลายชุดข้อมูลจะทำให้ผลการสืบค้นออกมาเป็นจำนวนมาก ทำให้ใช้เวลาในการประมวลผลสูงขึ้นมาก รวมถึงส่งผลต่อความแม่นยำในการสืบค้น (Precision) มีค่าความแม่นยำที่ต่ำ จากตัวอย่างได้ร้อยละ 59 เนื่องจากได้ผลการสืบค้นที่ไม่เกี่ยวข้องออกมาจำนวนมากตามไปด้วย

(2) การสืบค้นด้วยการวัดความคล้ายคลึงเชิงความหมาย (Semantic Similarity Score: SSS) การทำงานเป็นการสืบค้นโดยใช้การเปรียบเทียบคำสำคัญกับรายละเอียดภายในคลาสแต่ละคลาสที่มีค่าของคุณสมบัติประเภทอักขระ (Literal) หรือคำเหมือนเชิงความหมาย โดยใช้

กลุ่มคำศัพท์สัมพันธ์ในการเชื่อมโยงคำที่เกี่ยวข้องในเชิงลำดับชั้น (Broader Narrower) และคำที่มีความเกี่ยวข้องกันหรือใช้แทนกันได้ (Related, Alternative) ช่วยในการสืบค้น ซึ่งวิธีนี้จะทำให้ผลการสืบค้นมีความแม่นยำค่อนข้างสูง ได้ร้อยละ 97.82 เนื่องจากเอกสารที่ตรงกับคำสำคัญออกมาครบถ้วน แต่จะส่งผลต่อค่าความถูกต้องหรือค่าความระลึกได้ (Recall) ได้ร้อยละ78.73 เนื่องจากการสืบค้นด้วยคำสำคัญโดยไม่ดูกลุ่มคำศัพท์ของคุณสมบัติทรัพยากร (Predicate) จะทำให้การเชื่อมโยงไปยังเรื่องที่เกี่ยวข้องนั้นทำได้น้อยลงถ้าใช้คำศัพท์ที่ไม่ตรงหรือเกี่ยวข้องกับเรื่องที่สืบค้นจริง ๆ ทำให้ผลลัพธ์ที่ได้นั้นไม่ครบถ้วน แต่ทั้งนี้ถือว่าค่าที่ได้มีความแม่นยำและความถูกต้องในระดับยอมรับได้

(3) การรวมค่าคะแนนการวัดระยะทางเชิงความหมาย และวัดความคล้ายคลึงเชิงความหมายของคำสำคัญและคลาสในชุดข้อมูล โดยมีการให้น้ำหนักส่วนคะแนนทั้งสองเพื่อให้การสืบค้นมีมิติที่ครอบคลุมทั้งขอบเขตที่เหมาะสม และคำสำคัญที่มีความแม่นยำสูงเข้าด้วยกัน โดยมีการประเมินความเหมาะสมของค่าน้ำหนักตามโครงสร้างของชุดข้อมูลแต่ละชุดได้ ทำให้การสืบค้นมีความถูกต้องและแม่นยำสูงสุดเมื่อเทียบกับการใช้เพียงวิธีใดวิธีหนึ่ง เมื่อใช้วิธีการสืบค้นแล้วทำให้ค่าความแม่นยำได้ร้อยละ 95.55 ซึ่งมีความแม่นยำค่อนข้างสูง ถึงแม้จะไม่สูงเท่าวิธีการสืบค้นด้วยการวัดความคล้ายคลึงเชิงความหมายเพียงอย่างเดียวก็ตาม แต่ช่วยให้มีค่าความระลึกร้อยละ 88.91 ซึ่งเป็นค่าที่แสดงถึงรายการที่เกี่ยวข้องของเนื้อหามีความครบถ้วนสูงมาก เมื่อเปรียบเทียบประสิทธิภาพการสืบค้นทั้ง 3 วิธีแล้ว การใช้วิธีการสืบค้นโดยการรวมค่าคะแนนการวัดระยะทางเชิงความหมาย และการสืบค้นด้วยการวัดความคล้ายคลึงเชิงความหมายพร้อมกับการให้ค่าน้ำหนักอย่างเหมาะสม ช่วยให้การสืบค้นเลิร์นนิงอ็อบเจกต์ของชุดข้อมูลในโอเพนดาทามีประสิทธิภาพสูงขึ้นตามผลการทดลอง

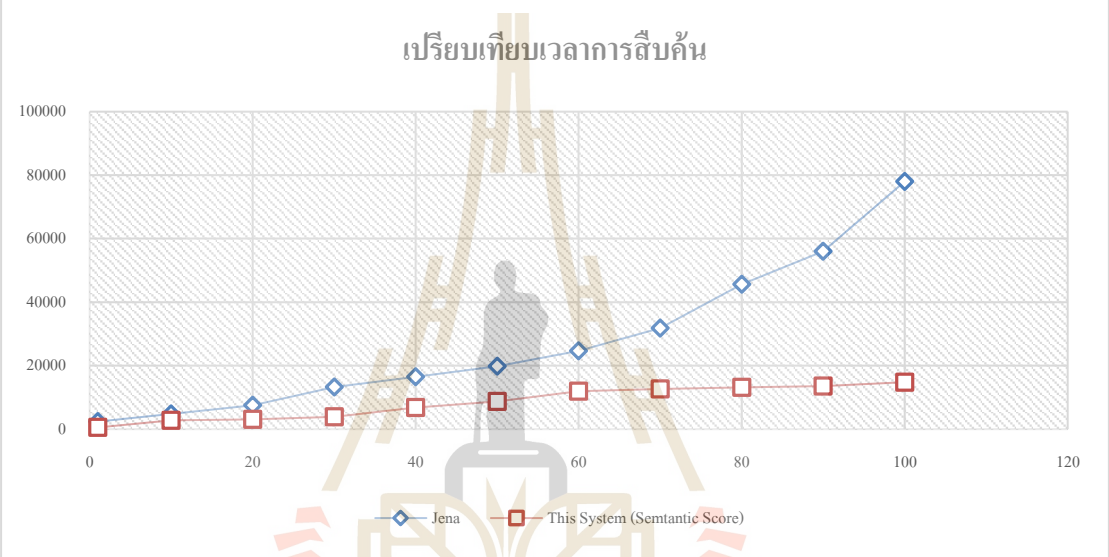
(2) การเปรียบเทียบประสิทธิภาพการทำงานของระบบกับระบบอื่น

ผู้วิจัยได้ทดสอบกับระบบจัดการอาร์ดีเอฟเจนา โดยมีรายละเอียดการทดลองดังนี้

(2.1) ชุดข้อมูลที่ใช้ในการทดสอบ ผู้วิจัยได้นำข้อมูลมูลตัวอย่างจากข้อมูลเลิร์นนิงอ็อบเจกต์ที่ใช้เป็นชุดข้อมูลในการทดสอบประสิทธิภาพระบบ โดยแบ่งชุดข้อมูลเป็น 5 ขนาดโดยมีขนาดที่แตกต่างกันตั้งแต่ 1MB 10MB 30MB 50MB 70MB และ 100MB โดยประมาณตามลำดับ ซึ่งแต่ละชุดข้อมูลมีจำนวนทริเปิลตั้งแต่พันระเบียนถึงหมื่นระเบียน เพื่อให้สะท้อนถึงขนาดชุดข้อมูลที่ใช้งานจริงในหลายขนาด

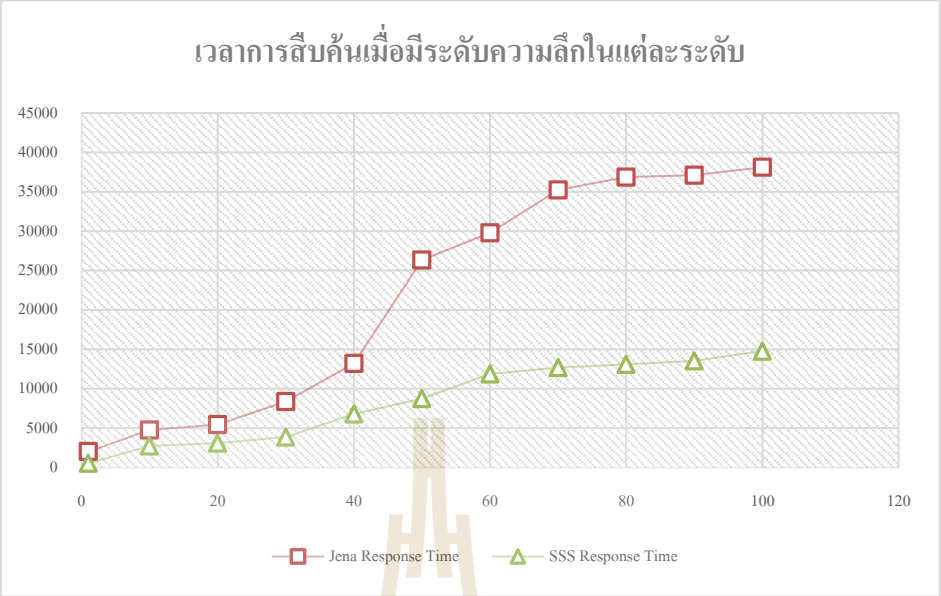
(2.2) ผลประสิทธิภาพการสืบค้น ผู้วิจัยได้ออกแบบการทดสอบโดยเปรียบเทียบกับการสืบค้นผ่านระบบจัดการข้อมูลอาร์ดีเอฟเจนา ซึ่งเป็นระบบแบบเปิดเผยรหัส โดยการทดลองนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของกระบวนการสืบค้น โดยการประเมิน

ถึงประสิทธิภาพของเวลาตอบสนอง (Response Time) ของระบบเจนาและระบบที่ออกแบบไว้ โดยการสืบค้นของระบบเจนาจะรองรับการสืบค้นด้วยภาษาสปาร์เคิล (SPARQL) และใช้การเปรียบเทียบแบบตรงคำค้น (Exact Match) เท่านั้น ส่วนระบบที่พัฒนาขึ้นจะทำการสืบค้นด้วยใช้การค้นเชิงความหมายโดยกำหนดค่า Threshold เป็น 0.5 และค่า p สำหรับการสืบค้นเท่ากับ 0.6 ตามผลการทดสอบประสิทธิภาพที่ได้จากกระบวนวัดประสิทธิภาพที่ผ่านมา โดยการทดสอบใช้คำในการค้น 1 คำ จำนวน 5 รอบเพื่อหาค่าเฉลี่ยเวลาที่ตอบสนอง โดยผลการทดสอบแสดงในรูปที่ 5.19



รูปที่ 5.19 เปรียบเทียบเวลาการสืบค้นระหว่างระบบเจนาและระบบที่พัฒนาขึ้นเอง

จากรูปที่ 5.19 เป็นกราฟแสดงเวลาการตอบสนองของระบบเจนาที่ระดับความลึกการสืบค้นเท่ากับ 1 เปรียบเทียบกับระบบที่พัฒนาขึ้น ซึ่งจะเห็นว่าเวลาในการสืบค้นระบบเจนาใช้เวลาการประมวลผลและสืบค้นสูงขึ้นตามขนาดของชุดข้อมูล เนื่องจากต้องสืบค้นข้อมูลทั้งชุดข้อมูล แต่ส่วนระบบที่พัฒนาขึ้นใช้เวลามากขึ้นเพียงเล็กน้อย เนื่องจากมีการคำนวณค่าระดับความลึกที่เหมาะสมและการเตรียมข้อมูลแบบตารางข้อมูลไว้ก่อนสามารถเข้าถึงข้อมูลจากดัชนีที่ทำไว้ส่งผลให้ขนาดข้อมูลที่มากขึ้นไม่มีผลต่อเวลาในกระบวนการค้นคืน ทั้งนี้ผู้วิจัยได้ทำการกำหนดระดับความลึกในการสืบค้นที่แตกต่างกันด้วยระบบที่พัฒนาขึ้น โดยให้มีความลึกในหลายระดับ พบว่าถึงแม้จะกำหนดระดับความลึกในหลายระดับ ความเร็วในการตอบสนองของระบบยังมีความรวดเร็วมากกว่าระบบเจนาที่ไม่มีการกำหนดระดับความลึก ดังแสดงในรูปที่ 5.20



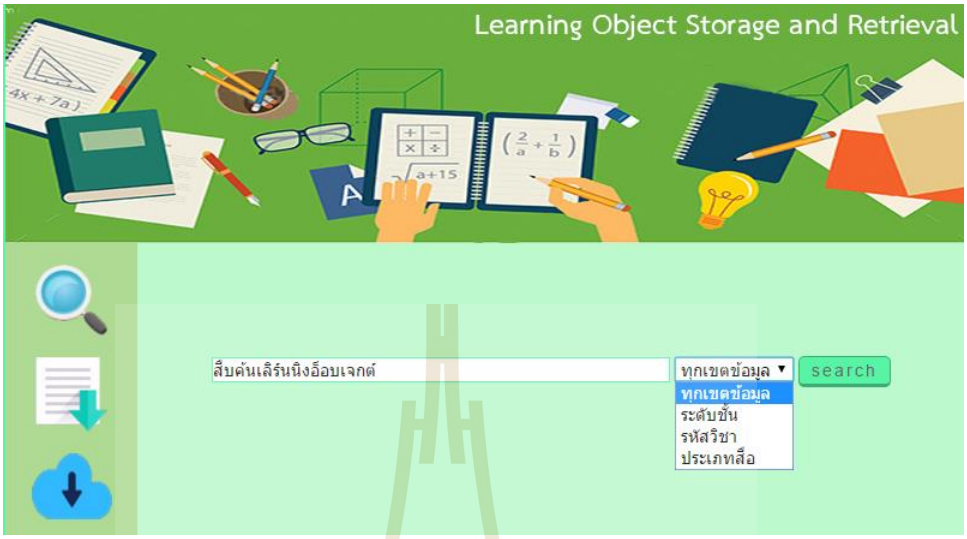
รูปที่ 5.20 เวลาการสืบค้นเมื่อมีการเพิ่มระดับความลึกในการสืบค้น

จากรูปที่ 5.20 แสดงเวลาการตอบสนองในระดับขนาดข้อมูลที่แตกต่างกัน โดยมีการกำหนดระดับความลึกที่สูงขึ้นในข้อมูลที่มีขนาดใหญ่ขึ้น ทั้งนี้จะเห็นว่าเมื่อเปรียบเทียบกับการประมวลผลของระบบเจนาในขนาดข้อมูลเท่ากัน เวลาการประมวลผลของระบบเจนาจะใช้เวลามากกว่าระบบที่พัฒนาขึ้น แสดงถึงการกำหนดระดับความลึกในการสืบค้นที่สูงขึ้น มีผลต่อเวลาการประมวลผลของระบบจริง แต่เนื่องด้วยระบบที่พัฒนาขึ้นมีการใช้แนวคิดการวัดความคล้ายคลึงมาช่วยในการพิจารณาโดยการเตรียมข้อมูลในแบบดัชนี (Index) ไว้ล่วงหน้าช่วยให้ลดระยะเวลาในการเปรียบเทียบการสืบค้นได้ ทำให้การสืบค้นด้วยระบบเจนาที่ใช้การสืบค้นแบบตรงคำค้น ต้องมีการประมวลผลการสืบค้นใหม่ทุกครั้งที่มีการสืบค้นทำให้ใช้เวลาการสืบค้นมากกว่า

5.3.2 การประเมินความสามารถของระบบโดยผู้ใช้

การประเมินความสามารถของระบบโดยผู้ใช้งานได้แก่ ผู้เชี่ยวชาญ และครูผู้สอน ผู้ใช้จะทำการทดสอบและประเมินระบบที่ติดตั้งไว้บนเครื่องแม่ข่ายจำลอง ผ่านเครื่องคอมพิวเตอร์ โดยผู้ใช้จะทำการประเมิน 2 ส่วนหลักได้แก่ (1) ส่วนการจัดเก็บเลิร์นนิงอ็อบเจกต์และการแปลงข้อมูล และ(2) ส่วนการสืบค้นเลิร์นนิงอ็อบเจกต์

หน้าจอการทำงานของระบบ มีรายละเอียดของระบบ ดังต่อไปนี้

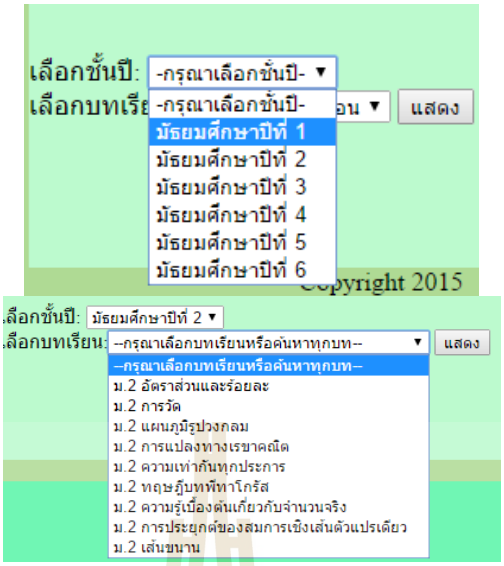


รูปที่ 5.21 หน้าหลักของระบบเมื่อเข้าใช้งาน

จากรูปที่ 5.21 หน้าหลักของระบบจะมีช่องสำหรับใส่คำค้น และทางด้านขวาจะมีปุ่มรายการเลือกเขตข้อมูล เพื่อสืบค้นเฉพาะบางเขตข้อมูลได้ ได้แก่ ค้นทุกเขตข้อมูล ค้นจากระดับชั้น ค้นจากรหัสวิชา และค้นจากประเภทสื่อ ซึ่งเขตข้อมูลเหล่านี้เป็นเขตข้อมูลที่ได้จากการสำรวจความนิยมในการเลือกเพื่อสืบค้นข้อมูลของครูผู้สอน ซึ่งตั้งเป็นค่าเริ่มต้นที่การสืบค้นทุกเขตข้อมูลเพื่อความสะดวกของผู้ใช้งานทั่วไป และมีปุ่ม Search เพื่อส่งคำค้นและเริ่มต้นการสืบค้น

ทางด้านซ้ายจะเป็นเมนูสำหรับการทำงานในคำสั่งการสืบค้นแบบไล่เรียง การนำข้อมูลเข้าจากภายนอก และการแปลงข้อมูลให้อยู่ในรูปแบบอาร์ดีเอฟ ตามลำดับ เพื่อเตรียมพร้อมการเชื่อมโยงไฟล์ข้อมูลขึ้นสู่โอเพนดาทาได้ทันที โดยมีรูปดังรูปที่ 5.22

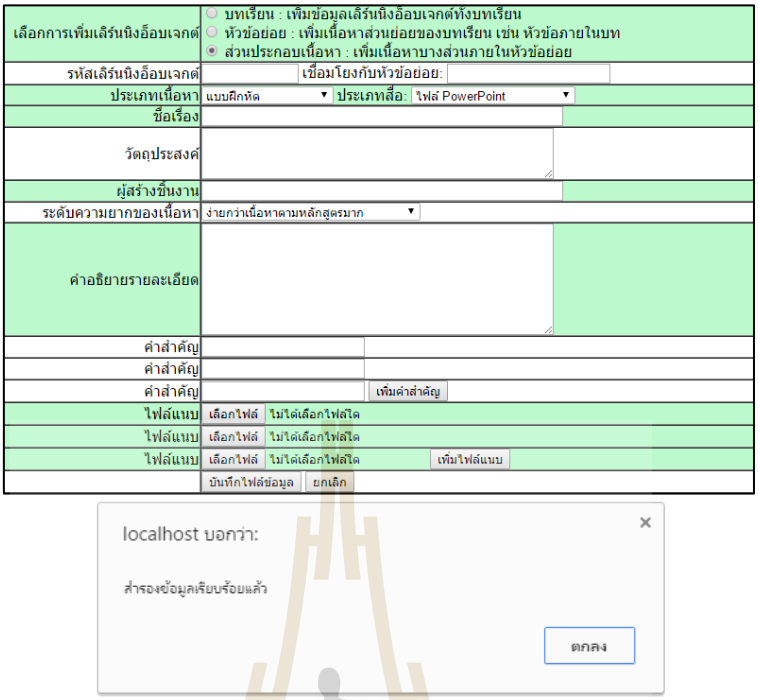
รูปที่ 5.22 เมนูย่อยเมื่อเข้าใช้งานคำสั่งการสืบค้นแบบไล่เรียง การนำเข้าข้อมูลและการแปลงไฟล์



รูปที่ 5.23 การสืบค้นแบบไล่เรียงตามระดับการศึกษาและบทเรียน

จากรูปที่ 5.23 เป็นการไล่เรียง (Browse) บทเรียนที่มีในระบบ โดยจะมีเมนูเลือกระดับชั้นเรียนที่ต้องการ จากนั้นระบบจะทำการแสดงชื่อบทเรียนในระดับชั้นนั้น ๆ ทุกบทเรียน โดยผู้ใช้สามารถเลือกบทเรียนที่ต้องการหรือเลือกแสดงเลิร์นนิงอ็อบเจกต์จากทุกบทเรียนในระดับชั้นได้ทันที โดยแสดงผลของเลิร์นนิงอ็อบเจกต์ ดังรูปที่ 5.24 แสดงบทเรียนทุกรายการในระดับชั้นมัธยมศึกษาปีที่ 2

รูปที่ 5.24 ผลการแสดงเลิร์นนิงอ็อบเจกต์แบบไล่เรียงตามชั้นปีการศึกษา

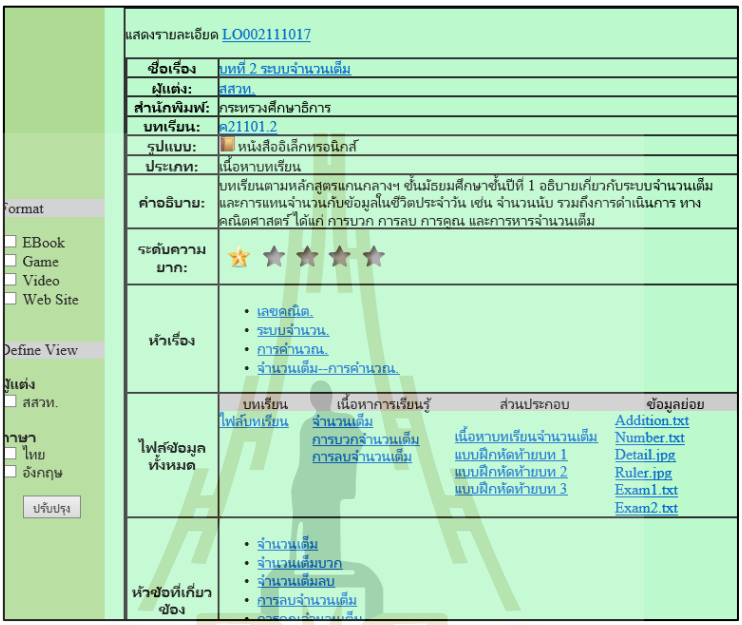


รูปที่ 5.25 หน้าจอเพิ่มเลิร์นนิงอ็อบเจกต์เข้าในระบบจัดเก็บเลิร์นนิงอ็อบเจกต์

จากรูปที่ 5.25 เป็นส่วนการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์โดยผู้ใช้ และเมื่อเพิ่มข้อมูลแล้วจะแสดงกล่องข้อความการสำรองข้อมูลเป็นอาร์ดีเอฟ (RDF) ไฟล์เรียบร้อย เนื่องจากเป็นกระบวนการแปลงชุดเมทาดาทาตามที่ออกแบบไว้แล้ว

รูปที่ 5.26 ผลการสืบค้นด้วยคำสำคัญโดยค้นทุกเขตข้อมูล

จากรูปที่ 5.26 เป็นการสืบค้นด้วยคำสำคัญ “จำนวนเต็ม” โดยค้นทุกเขตข้อมูล ซึ่งผลการสืบค้นจะแสดงเฉพาะระดับชั้น ชื่อเรื่อง และประเภทสื่อเป็นรายละเอียดเบื้องต้น ตามความต้องการของผู้ใช้จากแบบสำรวจ และมีข้อความ “แสดงรายละเอียด” เพื่อแสดงรายละเอียดของเลิร์นนิงอ็อบเจกต์เพิ่มเติม ดังภาพ 5.27



รูปที่ 5.27 ผลการแสดงรายละเอียดของเลิร์นนิงอ็อบเจกต์ที่เลือกจากรายการหลัก

จากรูปที่ 5.27 แสดงรายละเอียดของผลการสืบค้นเลิร์นนิงอ็อบเจกต์ ตามเขตข้อมูลต่าง ๆ โดยแสดงรายละเอียดของเมทาดาทาได้แก่ ชื่อเรื่อง ผู้แต่ง บทเรียน รูปแบบและประเภทสื่อ คำอธิบายเนื้อหา ระดับความยาก หัวเรื่อง ส่วนท้ายจะมีรายละเอียดขององค์ประกอบในเลิร์นนิงอ็อบเจกต์ทั้ง 4 ระดับ โดยในตัวอย่างแสดงรายละเอียดไฟล์ทั้งหมดในบทเรียนแบ่งออกเป็น 4 ส่วน ได้แก่ ไฟล์บทเรียนทั้งบท ไฟล์เนื้อหาการเรียนรู้แต่ละวัตถุประสงค์ในบทเรียน ไฟล์องค์ประกอบเนื้อหาย่อย เช่น ข้อความบทเรียน แบบฝึกหัด และส่วนข้อมูลย่อย ได้แก่ ไฟล์รูปภาพ ไฟล์ข้อความ

1) ผลการประเมินความสามารถของระบบโดยผู้ใช้

1.1) ข้อมูลทั่วไปของผู้ตอบแบบประเมินความสามารถระบบสืบค้นเลิร์นนิงอ็อบเจกต์ โดยกลุ่มผู้ใช้ที่เป็นผู้ประเมินระบบได้แก่ ครูผู้สอน และผู้เชี่ยวชาญ รวม 20 ท่านโดยรายละเอียดดังตารางที่ 5.19

ตารางที่ 5.19 ข้อมูลทั่วไปของผู้ตอบแบบประเมินความสามารถระบบสืบค้น

1. ข้อมูลทั่วไปของผู้ตอบแบบประเมิน	จำนวน	ร้อยละ
1. เพศ		
ชาย	4	20
หญิง	16	80
รวม	20	100
ข้อมูลทั่วไปของผู้ตอบแบบประเมิน	**จำนวน**	**ร้อยละ**
2. ระดับการศึกษา		
ต่ำกว่าปริญญาตรี	0	0
ปริญญาตรีหรือเทียบเท่า	14	70
ปริญญาโท	3	15
ปริญญาเอก	3	15
รวม	20	100

จากตารางที่ 5.19 พบว่า ผู้ตอบแบบประเมินความสามารถของระบบจัดเก็บและสืบค้นเลิร์น-นิงอ็อบเจกต์ ส่วนใหญ่เป็นเพศหญิง มีอัตราส่วนร้อยละ 80 ระดับการศึกษาปริญญาตรี ร้อยละ 70 รองลงมาคือ ระดับปริญญาโทและปริญญาเอก มีอัตราร้อยละเท่ากันที่ ร้อยละ 15

1.2) ความคิดเห็นเกี่ยวกับความสามารถของระบบ แบ่งเป็น 4 ด้าน ได้แก่ มอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์ มอดูลการนำเข้าข้อมูลจากภายนอกและการแปลงข้อมูล มอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์ และ 4 ความง่ายในการใช้งาน

ตารางที่ 5.20 ความคิดเห็นเกี่ยวกับความสามารถของระบบโดยรวมในแต่ละมอดูล

ความสามารถของระบบ	$\bar{X}$	S.D.	ระดับความคิดเห็น
1. มอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์	4.38	0.44	ดี
2. มอดูลการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์	4.87	0.34	ดีมาก
3. มอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์	4.37	0.52	ดี
4. ความง่ายในการใช้งาน	4.54	0.43	ดีมาก
โดยรวม	4.54	0.43	ดีมาก

มหาวิทยาลัยเทคโนโลยีสุรนารี

จากตารางที่ 5.20 พบว่า ผู้ประเมินระบบมีความเห็นว่าระบบโดยรวมมีประสิทธิภาพในระดับดีมาก ($\bar{X}$ = 4.54) เมื่อพิจารณาแยกเป็นองค์ประกอบย่อยทั้ง 4 ด้าน พบว่าองค์ประกอบที่ได้รับการประเมินในระดับ ดีมาก ได้แก่ มอดูลการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์ และเรื่องความง่ายในการใช้งานระบบโดยรวม ซึ่งมีค่าเฉลี่ย 4.87 และ 4.54 ตามลำดับ ส่วนองค์ประกอบที่ได้รับการประเมินในระดับ ดี ได้แก่ มอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์ และมอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์ ที่มีค่าเฉลี่ย 4.38 และ 4.37 ตามลำดับ

ตารางที่ 5.21 ความคิดเห็นเกี่ยวกับมอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์

1. มอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์	$\bar{X}$	S.D.	ระดับความคิดเห็น
1.1 เมนูการเพิ่มข้อมูลสื่อความหมาย และเข้าถึงได้ง่าย	4.89	0.32	ดีมาก
1.2 ฟอร์มกรอกข้อมูลมีเขตข้อมูลเหมาะสมชัดเจน	3.77	0.79	ปานกลาง
1.3 ลำดับการเพิ่มข้อมูลมีขั้นตอนที่เข้าใจง่าย	4.33	0.21	ดีมาก
1.4 ข้อมูลที่เพิ่มเสร็จมีความครบถ้วนตามต้นฉบับ	4.61	0.44	ดีมาก
โดยรวม	4.38	0.44	ดีมาก

จากตารางที่ 5.21 พบว่า ผู้ประเมินมีความเห็นเกี่ยวกับมอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์มีประสิทธิภาพในระดับดีมากอยู่ 3 ข้อ ได้แก่ เมนูการเพิ่มข้อมูลสื่อความหมายและเข้าถึงได้ง่าย ($\bar{X}$ = 4.89) ลำดับการเพิ่มข้อมูลมีขั้นตอนที่เข้าใจง่าย ($\bar{X}$ = 4.23) และข้อมูลที่เพิ่มเสร็จมีความครบถ้วนตามข้อมูลต้นฉบับ ($\bar{X}$ = 4.61) และประเมินองค์ประกอบในระดับปานกลาง 1 ข้อ ได้แก่ ฟอร์มการกรอกข้อมูลมีเขตข้อมูลที่เหมาะสมชัดเจน ($\bar{X}$ = 3.77)

ตารางที่ 5.22 ความคิดเห็นเกี่ยวกับมอดูลการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์จากภายนอก

2. มอดูลการนำเข้าข้อมูลจากภายนอก	$\bar{X}$	S.D.	ระดับความคิดเห็น
2.1 ความสะดวกในการเรียกใช้เมนู คำสั่ง	4.89	0.19	ดีมาก
2.2 กระบวนการนำเข้ามีขั้นตอนที่ง่าย ไม่ยุ่งยาก	5.00	0.48	ดีมาก
2.3 ระบบแจ้งขั้นตอนการดำเนินงานเข้าใจง่าย	4.81	0.11	ดีมาก
2.4 ข้อมูลมีความสมบูรณ์ครบถ้วนตามต้นฉบับ	4.76	0.59	ดีมาก
โดยรวม	4.87	0.34	ดีมาก

มหาวิทยาลัยเทคโนโลยีสุรนารี

จากตารางที่ 5.22 พบว่า ผู้ประเมิน เห็นว่าระบบมีประสิทธิภาพในระดับดีมาก ทุกประเด็น ได้แก่ ความสะดวกในการเรียกใช้งานเมนู และคำสั่ง ($\overline{X}$ = 4.89) กระบวนการนำเข้ามีขั้นตอนที่ง่าย ไม่ยุ่งยาก ($\overline{X}$ = 5.00) ระบบแจ้งขั้นตอนการดำเนินงานเข้าใจง่าย ($\overline{X}$ = 4.81) และข้อมูลที่จัดเก็บแล้วมีความสมบูรณ์ครบถ้วนตามข้อมูลต้นฉบับ ($\overline{X}$ = 4.76)

ตารางที่ 5.23 ความคิดเห็นเกี่ยวกับมอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์

3. มอดูลการค้นคืนเลิร์นนิงอ็อบเจกต์	$\overline{X}$	S.D.	ระดับความคิดเห็น
3.1 ช่องกรอกคำค้นมีตัวเลือกการสืบค้นเหมาะสม	4.56	0.32	ดีมาก
3.2 การสืบค้นแบบไร่เรียงมีเครื่องมือค้นที่เหมาะสม	4.78	0.18	ดีมาก
3.3 ผลการค้นหาแบบแบ่ง 2 ระดับมีความเหมาะสม	4.29	0.66	ดี
3.4 องค์ประกอบผลการสืบค้นเป็นประโยชน์ต่อการเลือกใช้	4.32	0.53	ดี
3.5 การจัดกลุ่มเนื้อหาที่เกี่ยวข้องมีความเหมาะสม	4.11	0.78	ดี
3.6 ระยะเวลาในการรอผลการสืบค้น	3.87	0.66	ปานกลาง
3.7 ผลการสืบค้นตรงกับความต้องการของผู้ใช้	4.69	0.51	ดีมาก
โดยรวม	4.37	0.52	ดี

จากตาราง 5.23 พบว่า ผู้ประเมินได้แสดงความคิดเห็นเกี่ยวกับประสิทธิภาพของมอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์ในระดับดีมาก จำนวน 3 ข้อ ได้แก่ ช่องกรอกคำค้นมีตัวเลือกการสืบค้นที่เหมาะสม ($\overline{X}$ = 4.56) การสืบค้นแบบไร่เรียง (Browse) มีเครื่องมือค้นที่เหมาะสม ($\overline{X}$ = 4.78) และ ผลการสืบค้นตรงกับความต้องการของผู้ใช้ ($\overline{X}$ = 4.69) และมีการประเมินองค์ประกอบในระดับ ดี จำนวน 3 ข้อ ได้แก่ ผลการค้นหาแบบแบ่ง 2 ระดับมีความเหมาะสม ($\overline{X}$ = 4.29) องค์ประกอบผลการสืบค้นเป็นประโยชน์ต่อการเลือกใช้ ($\overline{X}$ = 4.32) และ การจัดกลุ่มเนื้อหาที่เกี่ยวข้องมีความเหมาะสม ($\overline{X}$ = 4.11) และ ผลการประเมินในระดับ ปานกลาง มี 1 ข้อ ได้แก่ ระยะเวลาในการรอผลการสืบค้น ($\overline{X}$ = 3.87)

ตารางที่ 5.24 ความคิดเห็นเกี่ยวกับความง่ายในการใช้งานระบบจัดเก็บและสืบค้นเลิร์นนิงอ็อบ-เจกต์

ความง่ายในการใช้งานระบบจัดเก็บและสืบค้น	$\overline{X}$	S.D.	ระดับความคิดเห็น
1. ขนาด สี รูปแบบอักษร มีความเหมาะสม	4.88	0.21	ดีมาก

ตารางที่ 5.24 ความคิดเห็นเกี่ยวกับความง่ายในการใช้งานระบบจัดเก็บและสืบค้นเลิร์นนิงอ็อบ-เจกต์ (ต่อ)

ความง่ายในการใช้งานระบบจัดเก็บและสืบค้น	$\bar{X}$	S.D.	ระดับความคิดเห็น
2. การออกแบบกราฟิก มีความชัดเจน เหมาะสม	4.34	0.17	ดี
3. การจัดวางองค์ประกอบของระบบดูเหมาะสม	4.76	0.34	ดีมาก
4. แนวทางการเชื่อมโยงไปผลการสืบค้น และดาวน์โหลดเอกสาร มีความเหมาะสม	4.91	0.27	ดีมาก
5. รูปแบบของการระบุคำค้นสะดวกในการใช้งาน	4.25	0.68	ดี
6. ความเหมาะสมในการปฏิสัมพันธ์กับผู้ใช้	4.10	0.91	ดี
โดยรวม	4.54	0.43	ดีมาก

จากตารางที่ 5.24 พบว่า ผู้ประเมินมีความคิดเห็นต่อประสิทธิภาพของระบบในเรื่องความง่ายในการใช้งานระบบจัดเก็บและสืบค้นเลิร์นนิงอ็อบเจกต์ในระดับดีมาก มีจำนวน 3 ข้อ ได้แก่ ขนาดสี รูปแบบอักษร มีความเหมาะสม ($\bar{X}$ = 4.88) การจัดวางองค์ประกอบของระบบดูเหมาะสม ($\bar{X}$ = 4.76) และแนวทางการเชื่อมโยงไปผลการสืบค้น และดาวน์โหลดเอกสารแนบมีความเหมาะสม ($\bar{X}$ = 4.91) โดยประเมินความง่ายในการใช้งานระบบอยู่ในระดับดี จำนวน 3 ข้อ ได้แก่ การออกแบบกราฟิกมีความชัดเจน เหมาะสม ($\bar{X}$ = 4.34) รูปแบบของการระบุคำค้นสะดวกในการใช้งาน ($\bar{X}$ = 4.25) และความเหมาะสมในการที่ระบบมีปฏิสัมพันธ์กับผู้ใช้ ($\bar{X}$ = 4.10)

5.4 สรุปผลการออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา

ผลการประเมินประสิทธิภาพของระบบการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ที่สร้างจากโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทาที่พัฒนาขึ้น พบว่า ระบบที่พัฒนาขึ้นมีประสิทธิภาพในระดับดีมาก โดยมีค่าความแม่นยำ (Precision) ร้อยละ 95.55 ค่าการจำได้ (Recall) ร้อยละ 88.91 และค่าอัตราการรู้จำ (F-Measure) ร้อยละ 91.92 และผลการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญและครูผู้สอนพบว่า โดยรวมอยู่ในระดับ ดีมาก (ที่ค่าเฉลี่ย 4.54) ซึ่งแสดงให้เห็นถึงประสิทธิภาพของโมเดลที่ออกแบบและพัฒนาขึ้น

ในการออกแบบโมเดลระบบการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ด้วยโครงสร้างแบบเปิดเสรี ผู้วิจัยได้นำแนวคิดลิงค์โอเพนดาทาของเบอร์เนอร์สลี (Berners-Lee, 2006) มาเป็น

โครงสร้างหลักของไฟล์จัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์ โดยศึกษาถึงมาตรฐานและข้อกำหนดในการพัฒนาชุดข้อมูล (Dataset) ให้สอดคล้องกับแนวคิดการเปิดเสรีทางข้อมูล ได้แก่ การสร้างโครงสร้างไฟล์เป็นอาร์ดีเอฟ (RDF) และใช้ชุดคำศัพท์ควบคุมในการกำหนดชื่อคุณสมบัติของคลาสในระบบ ทั้งนี้ เพื่อสามารถเชื่อมโยงไปยังชุดข้อมูลอื่น ๆ ที่ใช้ชุดคำศัพท์ในกลุ่มเดียวกัน หรือเรื่องเดียวกันได้

นอกจากนี้ เพื่อให้ระบบสามารถรองรับข้อมูลเลิร์นนิงอ็อบเจกต์ที่ครูผู้สอนได้สร้างและจัดเก็บไว้ในรูปแบบอื่น ๆ เช่น ระบบฐานข้อมูลเชิงสัมพันธ์ ผู้วิจัยจึงได้ออกแบบขั้นตอนวิธีในการเปรียบเทียบและผสานเมทาดาทาจากระบบฐานข้อมูลเชิงสัมพันธ์ให้อยู่ในรูปแบบโครงสร้างอาร์ดีเอฟโดยมีเค้าร่างเมทาดาทาตามที่กำหนดในชุดข้อมูล เพื่อให้สามารถรวมข้อมูลเป็นชุดข้อมูลเดียวกันและสืบค้นได้พร้อมกัน โดยเครื่องมือที่ใช้ในการผสานคือ ชุดคำสั่งที่เปิดเผยรหัส D2RQ ที่ทำหน้าที่ในการผสานเค้าร่างเมทาดาทา ตามเงื่อนไขและข้อกำหนดในการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์ที่ผู้วิจัยกำหนดไว้ (Mapping Rule) โดยออกแบบตามการสำรวจความต้องการของผู้สอนในการใช้งานเค้าร่างเมาดาทาจากเขตข้อมูลต่าง ๆ

สำหรับมอดูลการสืบค้นเลิร์นนิงอ็อบเจกต์ ผู้วิจัยใช้วิธีการวัดความคล้ายคลึงเชิงความหมาย โดยมีส่วนประกอบย่อยในการวิเคราะห์ค่าคะแนนได้แก่ การวิเคราะห์คะแนนระยะทางเชิงความหมาย (Semantic Distance Score) และการวัดความคล้ายคลึงเชิงความหมาย (Semantic Similarity Score) ซึ่งผู้วิจัยได้ศึกษาถึงปัญหาการสืบค้นข้อมูลในลิงค์โอเพนดาทาในประเด็นปัญหาเกี่ยวกับ ความแม่นยำ และความถูกต้องของผลการสืบค้น กล่าวคือ การสืบค้นผ่านชุดข้อมูลนั้นผู้ออกแบบวิธีการสืบค้นจะต้องคำนึงถึงความสอดคล้องของคำค้นและข้อมูลในลิงค์โอเพนดาทาแต่ละคลาส และต้องกำหนดระดับความลึกของการสืบค้นหรือระดับการเชื่อมโยงผลการสืบค้นที่เกี่ยวข้องกันระหว่างชุดข้อมูลหรือคลาสต่าง ๆ เนื่องจากจะส่งผลต่อประสิทธิภาพของระบบ เพราะต้องสืบค้นข้อมูลที่มีปริมาณมากที่เชื่อมโยงถึงกันหมด ซึ่งส่งผลให้ประสิทธิภาพของการสืบค้นลดลงเพราะได้ผลการค้นที่เกินความจำเป็นและไม่เกี่ยวข้องกับความต้องการออกมามากเกินไป ดังนั้น ผู้วิจัยจึงใช้วิธีการคำนวณคะแนนทั้งสองส่วน โดยกำหนดค่าถ่วงน้ำหนักให้กับการวิเคราะห์ค่าคะแนนในด้านระยะทางเชิงความหมาย และค่าคะแนนความคล้ายคลึงเชิงความหมาย โดยสามารถปรับเปลี่ยนได้ตามโครงสร้างของชุดข้อมูล และความสัมพันธ์ภายในแต่ละคลาส ทั้งนี้ เพื่อให้ผลการสืบค้นมีความแม่นยำ และความถูกต้องสูงที่สุด และนอกจากการใช้ค่าถ่วงน้ำหนักแล้ว ยังทำการพัฒนาคลังศัพท์สัมพันธ์ด้านคณิตศาสตร์โดยใช้มาตรฐานสคอส (SKOS) เพื่ออธิบายความสัมพันธ์ของคำศัพท์ที่มีความสัมพันธ์แบบแคบกว่า (Narrower Term) คำศัพท์กว้างกว่า (Broader Term) และคำศัพท์ที่เกี่ยวข้องกัน (Related Term) เพื่อใช้ประกอบการวัดความคล้ายคลึง

ของคำค้นและคำสำคัญในคลาสที่อาจมีความเกี่ยวข้องสัมพันธ์กัน ทำให้ประสิทธิภาพการสืบค้นสูงยิ่งขึ้น

ในส่วนของการแสดงผล ผู้วิจัยออกแบบกระบวนการจัดลำดับผลการสืบค้น (Ranking) โดยใช้การวัดความคล้ายคลึงของคำค้น และผลการสืบค้นด้วยแนวคิดปริภูมิเวกเตอร์ (Vector Space) ด้วยสมการแบบโคซายน์ (Cosine) และประยุกต์สูตรโดยเพิ่มค่าน้ำหนักของคำสำคัญตามปัจจัยที่ต้องการจัดลำดับ

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 6
สรุปผลการวิจัยและข้อเสนอแนะในการทำวิจัย

บทนี้จะนำเสนอเกี่ยวกับสรุปผลการวิจัย การประยุกต์ใช้ผลการวิจัย และข้อเสนอแนะในการวิจัยครั้งต่อไป โดยมีรายละเอียด ดังนี้

6.1 สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา โดยอาศัยแนวคิดทฤษฎี และเทคนิควิธีต่าง ๆ ซึ่งกระบวนการวิจัยสามารถแบ่งออกได้ 3 ส่วน ประกอบด้วย (1) การศึกษากระบวนการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ (2) การออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ และ (3) การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ โดยผลการวิจัยในแต่ละขั้นตอนสามารถสรุปรายละเอียดได้ ดังต่อไปนี้

6.1.1 การศึกษากระบวนการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

การศึกษากระบวนการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ แบ่งกระบวนการศึกษาออกเป็น 2 ส่วน ได้แก่ (1) การศึกษาพฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอนจากโรงเรียนในสังกัดของ สพฐ. และโรงเรียนเอกชนที่มีการเรียนการสอนในหลักสูตรสองภาษา และ (2) การศึกษาเค้าร่างเมทาดาทาที่ใช้ในระดับสากลได้แก่ มาตรฐานดับลินคอร์ และมาตรฐานลอม ผลการวิจัย พบว่า

1) พฤติกรรมการค้นคืนเลิร์นนิงอ็อบเจกต์ของครูผู้สอน

1.1) แหล่งข้อมูลที่ครูผู้สอนใช้ในการค้นคืนเลิร์นนิงอ็อบเจกต์ ผลการศึกษาพบว่า ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษาใช้หนังสือเรียนตามหลักสูตรแกนกลางของสพฐ. เป็นแหล่งข้อมูลหลักในการเตรียมเอกสารการเรียนการสอนมากที่สุดและใช้เสิร์ชเอ็นจินสืบค้นข้อมูลเพิ่มเติมจากอินเทอร์เน็ต เนื่องจากสามารถเข้าถึงข้อมูลจากแหล่งอื่นได้รวดเร็ว รวมถึงได้ข้อมูลในรูปแบบดิจิทัลที่สามารถนำมาใช้งานได้สะดวก โดยนิยมนำสื่อการเรียนที่ผู้สร้างไว้เรียบร้อยแล้วมาใช้มากกว่าการนำองค์ประกอบมาสร้างเป็นสื่อการเรียนใหม่ เนื่องจากปัญหาด้านลิขสิทธิ์ อีกทั้งมีความยุ่งยากกว่าการนำสื่อสำเร็จรูปมาใช้

1.2) หนังสือเรียนเสริม และสื่อประสมนั้น ยังคงมีข้อจำกัดด้านงบประมาณในการจัดหา นอกจากนี้ยังพบว่า ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษา ยังคงนิยมใช้ข้อมูลจากแหล่งข้อมูลที่น่าเชื่อถือ คือ หน่วยงานทางการศึกษาของภาครัฐ อาทิ กระทรวงศึกษาธิการ สำนักงานส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี เป็นต้น เนื่องจากเป็นหน่วยงานที่รับผิดชอบหลักสูตรโดยตรง ทำให้สามารถหาเลิร์นนิงอ็อบเจกต์ได้ตรงตามเนื้อหาของหลักสูตรการเรียนการสอนมากที่สุด อย่างไรก็ตาม จากการศึกษาพบว่า เว็บไซต์ของหน่วยงานดังกล่าวยังมีการจัดทำและสร้างเลิร์นนิงอ็อบเจกต์เพื่อเผยแพร่ในจำนวนน้อยมากและมีเพียงบางวิชา หรือบาง หัวเรื่องเท่านั้น

1.3) วิธีการสืบคืนและคัดเลือกรายการเลิร์นนิงอ็อบเจกต์บนอินเทอร์เน็ต พบว่า วิธีการสืบคืนที่ครูผู้สอนใช้มากที่สุดคือ คำสำคัญที่เกี่ยวข้องกับเรื่องที่ต้องการสืบค้น โดยมักใช้คำค้นที่เป็นคำสำคัญในเนื้อหา พอ ๆ กับชื่อเรื่องของบทเรียนเนื่องจากเชื่อว่าจะได้ผลการค้นคืนที่เกี่ยวข้องกับเรื่องที่ต้องการมากที่สุด ส่วนข้อมูลที่ไม่นิยมใช้ในการสืบคืนคือ ชื่อหนังสือ และชื่อบุคคลที่เกี่ยวข้องกับเนื้อหานั้น เพราะจะได้ข้อมูลที่กว้างเกินไป ส่วนการสืบค้นจากข้อมูลหลายตัวหรือการค้นขั้นสูง (Advance Search) ไม่เป็นที่นิยม เนื่องจากผลการสืบคืนจะเฉพาะเจาะจงมากเกินไป ซึ่งอาจจะไม่มีรายการที่ตรงกับเงื่อนไขออกมาเลย และการกำหนดเงื่อนไขจำนวนมากจะทำให้ผลการสืบค้นจากเรื่องอื่น ๆ ที่อาจเกี่ยวข้องถูกตัดออกไปด้วย

สำหรับเนื้อหาที่ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษานิยมสืบค้น ได้แก่ บทบรรยายและเนื้อหาการสอนหลัก เนื่องจาก สามารถนำข้อความเนื้อหาที่เป็นแบบดิจิทัลไปใช้งานได้ทันที นอกจากนี้ ยังนิยมค้นหาแบบฝึกหัด ข้อสอบ และโจทย์ปัญหาเพื่อนำไปใช้เป็นแบบฝึกหัดที่มีความหลากหลาย และสามารถดัดแปลงให้มีความซับซ้อน แปลกใหม่ได้ง่าย โดยไม่ต้องเริ่มออกข้อสอบเองตั้งแต่ต้น ส่วนภาพประกอบ เกม และกิจกรรม ไม่เป็นที่นิยมในการสืบค้น เนื่องจาก ประเด็นปัญหาด้านลิขสิทธิ์และการนำมาดัดแปลงเพื่อใช้งานนั้นทำได้ยาก

1.4) การนำเลิร์นนิงอ็อบเจกต์ที่ได้ไปใช้ในการเรียนการสอน พบว่า เลิร์นนิงอ็อบเจกต์ที่ครูผู้สอนค้นคืนไปนั้น ส่วนใหญ่ถูกนำไปใช้เป็นส่วนประกอบของเนื้อหาการสอนแบบเสริม หรือ เนื้อหาที่เพิ่มเติมจากเอกสารประกอบการสอนหลัก เช่น แบบฝึกหัดเพิ่มเติม แบบทดสอบนอกบทเรียนเป็นต้น โดยส่วนเสริมเหล่านี้ครูผู้สอนต้องเป็นผู้สร้างขึ้นเอง เพิ่มเติมจากหนังสือเรียนตามหลักสูตรแกนกลาง นอกจากนี้ ยังมีการนำไปใช้ในการสอนหลักหากเนื้อหาหลักมีรายละเอียดไม่เพียงพอ หรือในกรณีที่ครูผู้สอนต้องการตัวอย่างที่แตกต่างจากตัวอย่างในหนังสือเรียน ส่วนการเขียนตำราหรือเอกสารการสอนนั้นครูผู้สอนจะนิยมอ้างอิงจากเนื้อหาหนังสือเรียนตามหลักสูตรมากกว่าเนื่องจากเห็นว่าเนื้อหามีการเผยแพร่อย่างเป็นทางการ

1.5) ปัญหาที่พบในการสืบค้นเลิร์นนิงอ็อบเจกต์ ผลการศึกษาพบว่า ครูผู้สอนคณิตศาสตร์ระดับมัธยมศึกษาหลักสูตรสองภาษามักจะมีปัญหากับการเลือกใช้คำค้นที่เป็นภาษาอังกฤษ เนื่องจากมักทำให้ได้ผลการค้นเป็นภาษาอังกฤษ และการพิจารณาเลือกผลการค้นที่เป็นภาษาอังกฤษยังทำได้ยากเนื่องจากไม่แน่ใจในรายละเอียดของเนื้อหา ปัญหาด้านการนำไปใช้ พบว่า ประเด็นที่ผู้ตอบแบบสอบถามเห็นว่าเป็นปัญหามากที่สุดคือ ข้อมูลที่ต้องการมีค่าใช้จ่ายหรือมีลิขสิทธิ์ ซึ่งทำให้ไม่สามารถนำไปใช้ได้อย่างเสรี หรือถูกปิดกั้นการเข้าถึง รองลงมาคือ รายการผลการค้นมีความซ้ำซ้อนหรือมีความเหมือนกัน หรือ ผลการสืบค้นมีหลายรายการแต่เชื่อมโยงไปยังจุดหมายเดียวกัน ซึ่งเป็นลักษณะของผลการสืบค้นจากเสิร์ชเอ็นจิน ทำให้ผู้ค้นเกิดความสับสนและใช้เวลาสืบค้นหน้าเอกสารใหม่ ๆ นานขึ้น

1.6) การจัดเก็บเลิร์นนิงอ็อบเจกต์ พบว่า ครูผู้สอนส่วนใหญ่จัดเก็บไฟล์ของเลิร์นนิงอ็อบเจกต์ไว้ในเครื่องคอมพิวเตอร์ของตนเอง และบางส่วนจะไม่เก็บไฟล์อิเล็กทรอนิกส์แต่จะพิมพ์ออกมาเป็นเอกสารกระดาษและจัดเก็บไว้ และทำการสร้างใหม่ทุกครั้งที่ต้องการใช้งาน ส่วนการใช้ระบบจัดเก็บข้อมูลของส่วนกลางนั้นมีจำนวนน้อย เนื่องจากโรงเรียนไม่ได้จัดหาระบบส่วนกลางให้เป็นทางการ

2) การออกแบบชุดเมทาดาทาสำหรับอธิบายเลิร์นนิงอ็อบเจกต์

ชุดเมทาดาทาที่ใช้อธิบายเลิร์นนิงอ็อบเจกต์ในระบบจัดเก็บเลิร์นนิงอ็อบเจกต์ในงานวิจัยนี้ ออกแบบและพัฒนาขึ้นจากข้อมูล 2 ส่วน ได้แก่

2.1) เมทาดาทาที่ได้จากการศึกษาพฤติกรรมการสืบค้นเลิร์นนิงอ็อบเจกต์ของครูผู้สอน พบเมทาดาทาที่ผู้ใช้เลือกใช้ในการพิจารณารายละเอียดของเลิร์นนิงอ็อบเจกต์ในระดับ “มาก” ขึ้นไปรวมจำนวน 10 องค์ประกอบ ได้แก่ ชื่อเรื่อง (Title) คำสำคัญ (Keyword) ประเภทของไฟล์ (Format) รายละเอียด (Description) บทคัดย่อ (Abstract) หัวเรื่อง (Subject) ภาษา (Language) แหล่งที่มาของข้อมูล (Source) ระดับความยากง่ายของเนื้อหา (Difficulty) ผู้รับผิดชอบเนื้อหา (Publisher)

2.2) เมทาดาทาที่ศึกษาจากงานวิจัยและข้อกำหนดตามมาตรฐานโครงสร้างด้านกายภาพและเนื้อหาของเลิร์นนิงอ็อบเจกต์ที่ควรมี ได้องค์ประกอบเมทาดาทาเพิ่มเติมในชุดโครงสร้างอีก 7 องค์ประกอบ ได้แก่ ระดับการศึกษาหรือกลุ่มเป้าหมาย (EducationLevel/Audience) รูปแบบการปฏิสัมพันธ์ (Interactive Type) ระดับการปฏิสัมพันธ์ (Interactive Level) ประเภทเลิร์นนิงอ็อบเจกต์ (Learning object Type) ระยะเวลาการเรียน (Learning Time) มีส่วนประกอบ (hasPart) เป็นส่วนประกอบของ (isPartOf)

เมื่อรวบรวมเมทาดาทาจากทั้ง 2 ส่วนจะได้เป็นชุดเมทาดาทาที่ใช้อธิบายเลิร์นนิงอ็อบเจกต์ในงานวิจัยนี้ทั้งสิ้น 17 องค์ประกอบ โดยนำองค์ประกอบย่อยจากมาตรฐานเมทาดาทาที่นิยมใช้ในการอธิบายเลิร์นนิงอ็อบเจกต์ ได้แก่ มาตรฐาน Dublin Core มาตรฐาน และ IEEE LOM มาปรับใช้กับองค์ประกอบย่อยเมทาดาทาที่ได้จากการศึกษาครั้งนี้

6.1.2 ผลการออกแบบและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

ผลการออกแบบและพัฒนาโมเดลจะแบ่งงานออกเป็น 2 ส่วนได้แก่ (1) การออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ และ (2) การออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ โดยมีรายละเอียดดังต่อไปนี้

1) การออกแบบและพัฒนาโมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์ มีองค์ประกอบทั้งหมด 3 ส่วนได้แก่

1.1) ระบบคลังข้อมูลเลิร์นนิงอ็อบเจกต์แบบฐานข้อมูลเชิงสัมพันธ์ เพื่อรองรับการเพิ่มข้อมูลเลิร์นนิงอ็อบเจกต์ที่สร้างขึ้นใหม่ หรือเป็นการนำรายการเลิร์นนิงอ็อบเจกต์เข้ามาจากภายนอกระบบ โดยนำชุดขององค์ประกอบย่อยเมทาดาทาที่พัฒนาขึ้นมาใช้บรรยายคุณลักษณะของเลิร์นนิงอ็อบเจกต์ที่จะจัดเก็บในคลังข้อมูล ซึ่งงานวิจัยนี้ได้แบ่งคลาสของข้อมูลภายในคลังจัดเก็บเลิร์นนิงอ็อบเจกต์ ออกเป็น 10 คลาส ได้แก่ ผู้สร้างผลงาน (Contributor) แผนกและส่วนงาน (Department) หน่วยบทเรียน (Lesson) วัตถุประยุกต์ (Application Object) วัตถุสารสนเทศ (Information Object) สื่อข้อมูลย่อย (Asset) คลาสคำสำคัญ (Keyword) คลาสรูปแบบการมีปฏิสัมพันธ์ (InteractiveType) คลาสประเภทเลิร์นนิงอ็อบเจกต์ (ResourceType) และคลาสภาษา (Languages) ตัวอย่างทดลองการวิจัยนี้ได้รวบรวมจากเว็บไซต์ที่เผยแพร่เลิร์นนิงอ็อบเจกต์ในรายวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 1-6 ซึ่งเป็นระบบแบบเปิดเสรีทั้งในประเทศและต่างประเทศ จำนวน 300 ระเบียน

1.2) การเปรียบเทียบและผสานเมทาดาทาของเลิร์นนิงอ็อบเจกต์ที่รวบรวมจากคลังข้อมูลที่ใช้เค้าร่างเมทาดาทาต่างกันให้จัดเก็บตามเมทาดาทาที่ได้ออกแบบไว้ในข้อ (1.1) และแปลงข้อมูลให้อยู่ในโครงสร้างแบบอาร์ดีเอฟ (RDF) ตามมาตรฐานลิงค์โอเพนดาทา ซึ่งการพิจารณาความสัมพันธ์ของฐานข้อมูลเชิงสัมพันธ์มี 2 รูปแบบ ดังนี้

1.2.1) กรณีความสัมพันธ์แบบหลายต่อหลาย (Many to Many: M:M) จะเกิดอ็อบเจกต์พร็อพเพอร์ตี้ระหว่างสองตาราง โดยตารางทั้งสองจะถูกสร้างเป็นคลาส และคุณสมบัติระหว่างคลาสจะเป็นส่วนกลับความสัมพันธ์ (inversProperty)

1.2.2) กรณีความสัมพันธ์แบบหนึ่งต่อหลาย (One to Many: 1:M) ความสัมพันธ์ของตารางในรูปแบบนี้สามารถแปลงได้ 2 ลักษณะ คือ (1) กรณีที่ตารางทั้งสองไม่ใช่

ข้อมูลที่เป็นซับเซตกัน หรือเป็นข้อมูลคนละประเภท จะทำการสร้างอ็อบเจกต์พร็อพเพอร์ตี้ระหว่างคลาสทั้งสอง โดยกำหนดให้เรนจ์ (Range) ของอ็อบเจกต์พร็อพเพอร์ตี้อยู่ทางคลาสที่มาจากตารางที่มีความสัมพันธ์แบบหนึ่ง ส่วนโดเมน (Domain) จะอยู่ทางคลาสที่มาจากตารางที่มีความสัมพันธ์แบบหลาย และ (2) กรณีที่ตารางทั้งสองเป็นข้อมูลที่เป็นซับเซตกัน หรือมีข้อมูลในกลุ่มเดียวกัน จะทำการกำหนดความเป็นคลาสหลักและคลาสย่อย (Class and SubClass) ระหว่างตารางทั้งสองเมื่อเปลี่ยนตารางเป็นคลาส คลาสที่มีความสัมพันธ์แบบหลาย (Many) จะเป็นคลาสย่อย (SubClass) ของคลาสที่มีความสันพันธ์แบบหนึ่ง (One)

คลาสที่ได้จากการแปลงฐานข้อมูลได้แก่ คลาสผู้สร้างผลงาน (Contributor) มีความสัมพันธ์กับแผนกหรือส่วนงาน (Department) ผ่านอ็อบเจกต์พร็อบเพอร์ตี้ workInDepartment ผู้สร้างผลงานสัมพันธ์กับรายวิชา (Course) ผ่านอ็อบเจกต์พร็อพเพอร์ตี้ isCreateCourse สัมพันธ์กับบทเรียน (Lesson) ผ่านอ็อบเจกต์พร็อพเพอร์ตี้ isCreateLesson และสัมพันธ์กับเลิร์นนิงอ็อบเจกต์ (Learning Object) และส่วนประกอบเนื้อหา (Asset) ผ่านอ็อบเจกต์พร็อพเพอร์ตี้ isCreateContentObject และ isCreateAsset ตามลำดับ

คลาสรายวิชา (Course) สัมพันธ์กับคลาสบทเรียน (Lesson) ผ่านอ็อบเจกต์พร็อพเพอร์ตี้ consistOfLesson เพื่อเชื่อมโยงบทเรียนแต่ละบทเข้ากับรายวิชา

คลาสบทเรียน (Lesson) สัมพันธ์กับคลาสย่อยบท (Chapter) และคลาสย่อยส่วน (Section) ผ่านอ็อบเจกต์พร็อพเพอร์ตี้ hasChapter และ hasSection ตามลำดับ โดยทั้งสองคลาสนี้จะใช้เพื่อเก็บรายละเอียดย่อยในกรณีที่บทเรียนมีเนื้อหาแยกเป็นบทย่อย หรือเป็นส่วน เพื่อความสะดวกในการค้นหาหรือจัดการเนื้อหาบางส่วน และยังสัมพันธ์กับคลาสเลิร์นนิงอ็อบเจกต์ (LearniungObject) ผ่านอ็อบเจกต์พร็อพเพอร์ตี้ hasContentObject เพื่อกำหนดความสัมพันธ์ระหว่างบทเรียนแต่ละบทมีส่วนประกอบของเลิร์นนิงอ็อบเจกต์ร่วมอยู่ได้หลายรายการ
คลาสเลิร์นนิงอ็อบเจกต์หรือคอนเท็นอ็อบเจกต์ (Content Object) มีความสัมพันธ์กับคลาสส่วนประกอบเนื้อหา (Asset) ผ่านอ็อบเจกต์พร็อพเพอร์ตี้ consistOfAsset เพื่อเชื่อมโยงส่วนประกอบเนื้อหาที่นำมาเป็นส่วนประกอบในการจัดทำเลิร์นนิงอ็อบเจกต์

1.3) การสร้างคลังคำศัพท์รายวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษาปีที่ 1 ถึง 6 เพื่อนำมาใช้ในการจัดเก็บและค้นคืนเลิร์นนิ่งอ็อบเจ็กต์ โดยจัดเก็บคำศัพท์ตามโครงสร้างความสัมพันธ์ของสคอส (Simple Knowledge Organization System: SKOS) ซึ่งมีการกำหนดความสัมพันธ์ของกลุ่มคำศัพท์ใน 3 กลุ่ม ได้แก่ (1) คำศัพท์ที่มีความสัมพันธ์ในลักษณะที่เป็นทางเลือกหรือใช้ทดแทนกันได้ (Alternative Term) (2) คำศัพท์ที่มีความสัมพันธ์ในลักษณะลดหลั่นในระดับชั้นที่กว้างกว่าและแคบกว่ากว่า (Broader Term, Narrower Term) และ (3) ความสัมพันธ์ใน

ลักษณะที่เกี่ยวข้องกัน หรือสัมพันธ์กันในบางประเด็น (Related Term) จากนั้นจัดเรียงคำศัพท์ตามโครงสร้างความสัมพันธ์ภายใต้หัวเรื่องคณิตศาสตร์ของเวิร์ดเน็ต ซึ่งคำศัพท์ที่รวบรวมมาจัดเก็บไว้ในงานวิจัยนี้มีทั้งสิ้น 380 คำโดยแหล่งรวบรวมคำศัพท์จาก 4 แหล่ง ได้แก่ (1) หนังสือเรียนรายวิชาพื้นฐานคณิตศาสตร์ และหนังสือเรียนรายวิชาเพิ่มเติมคณิตศาสตร์ ระดับชั้นมัธยมศึกษาชั้นปีที่ 1 ถึง 6 ตามการกำหนดของหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน จำนวน 16 เล่ม (2) พจนานุกรมคำศัพท์คณิตศาสตร์สำหรับนักเรียน (3) คำศัพท์สัมพันธ์จากเวิร์ดเน็ตจากคำหลักคณิตศาสตร์และคำที่เกี่ยวข้อง และ (4) คำศัพท์จาก เลิร์นนิงอ็อบเจกต์ จำนวน 300 รายการ ทำการประเมินประสิทธิภาพของคลังคำศัพท์โดยผู้เชี่ยวชาญจำนวน 5 คน และครูผู้สอนรายวิชาคณิตศาสตร์ระดับมัธยมศึกษาจำนวน 5 คน ผลการประเมินพบว่าคลังคำทางคณิตศาสตร์ที่รวบรวมขึ้นมามีคำศัพท์ที่ครอบคลุมหัวเรื่องของบทเรียนคณิตศาสตร์ระดับมัธยมศึกษาในระดับดีมาก มีความสอดคล้องกับคำศัพท์ภายในเนื้อหาบทเรียนตามหลักสูตรแกนกลางวิชาคณิตศาสตร์ตามที่กระทรวงศึกษาธิการกำหนดไว้ นอกจากนี้การกำหนดความสัมพันธ์ของคำศัพท์มีความสามารถในการแนะนำคำศัพท์ที่เกี่ยวข้องทั้งตามเนื้อหาและตามระดับชั้นเรียนได้ สรุปความพึงพอใจของผู้ประเมินคลังมีความพึงพอใจในระดับมากที่สุด

2) การออกแบบและพัฒนาโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ มีองค์ประกอบทั้งหมด 2 ส่วนได้แก่

2.1) การวัดความคล้ายคลึงเชิงความหมาย เป็นการออกแบบเพื่อนำเข้ามาช่วยในการเปรียบเทียบคำค้นจากชุดข้อมูลที่ได้จากลิงค์โอเพนดาทา โดยใช้การคำนวณจากค่าคะแนนระยะทางเชิงความหมาย (Semantic Distance Score) และ การวัดความคล้ายคลึงเชิงความหมาย (Semantic Similarity Score) โดยการปรับแต่งโครงสร้างชุดข้อมูลให้เป็นโครงสร้างดัชนีแบบจุดยอด (Vertex) ตามโหนดในโครงสร้างอาร์ดีเอฟ ซึ่งจะเก็บข้อมูลเป็นคู่อ็อบเจ็กต์ที่ส่งเข้า (Incoming Node) และนำออก (Outgoing Node) โดยการวัดความคล้ายคลึงเชิงความหมายจะมีค่าถ่วงน้ำหนักให้กับค่าคะแนนทั้งสอง ทั้งนี้เพื่อให้การเปรียบเทียบคำค้นกับคำสำคัญสามารถเข้าถึงโหนดที่เกี่ยวข้องตามคุณสมบัติของทรัพยากรที่เกี่ยวข้อง (Predicate) และโหนดที่มีข้อมูลหรือความคล้ายคลึงกับคำค้นได้อย่างเหมาะสม ซึ่งผลการทดลองพบว่า การใช้ค่าพารามิเตอร์ p มาช่วยในการปรับค่าน้ำหนักของการวัดค่าระยะทางเชิงความหมาย และความคล้ายคลึงเชิงความหมาย มีผลต่อความแม่นยำ ความถูกต้อง และอัตราการรู้จำที่จะช่วยให้มีประสิทธิภาพสูงขึ้น โดยมีค่าความแม่นยำ (Precision) ร้อยละ 95.55 ค่าความถูกต้องหรือการจำได้ (Recall) ร้อยละ 88.91 และค่าอัตราการรู้จำ (F-Measure) ร้อยละ 91.92

2.2) การจัดลำดับการนำเสนอผลการค้นคืน ในงานวิจัยนี้ประกอบด้วยแนวทางการจัดลำดับผลการค้นคืนด้วยวิธีการคิวรี่ดีเพนเด็นท์แรงกิ้ง (Query Dependent Ranking) (Liu, 2009) ซึ่งเป็นการจัดลำดับผลการค้นคืนโดยยึดการเปรียบเทียบระหว่างคำค้นกับคำดัชนีตัวแทนของเอกสารเป็นสำคัญ นอกจากนี้ ยังได้เพิ่มแนวคิดการเพิ่มค่าน้ำหนักของคำค้นและองค์ประกอบทางบรรณานุกรม (Bibliographic) เข้าไปเพื่อเปรียบเทียบผมการจัดลำดับที่ได้กับการจัดลำดับโดยใช้ข้อมูลตัวแทนเอกสารแบบไม่มีการใช้ค่าน้ำหนัก นำมาเข้าสมการคำนวณหาความคล้ายคลึงระหว่างข้อมูลสองรายการแบบปริภูมิเวกเตอร์ (Vector Space) ด้วยวิธีการวัดความคล้ายคลึงแบบโคซายน์ (Cosine Similarity) และได้ทำการปรับปรุงให้รองรับการเพิ่มค่าน้ำหนักของข้อมูลจากบรรณานุกรม (เช่น ชื่อผู้แต่ง หน่วยงาน) ที่ต้องการในการวัดความคล้ายคลึงได้ด้วย

6.1.3 การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

การประเมินโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลโอเพนดาทา ได้กำหนดแนวทางการประเมินไว้ 2 วิธี ได้แก่ (1) การประเมินประสิทธิภาพของระบบที่พัฒนาขึ้นตามโมเดลด้วยการเปรียบเทียบความเร็วในการค้นคืนเลิร์นนิงอ็อบเจกต์กับเจนาเฟรมเวิร์ค (Jena Framework) (2) การประเมินการใช้งานระบบที่พัฒนาขึ้นตามโมเดลโดยผู้ใช้

1) การประเมินประสิทธิภาพของระบบที่พัฒนาขึ้นตามโมเดลด้วยการเปรียบเทียบการทำงานกับเจนาเฟรมเวิร์ค ซึ่งเป็นชุดคำสั่งสำเร็จรูปในการค้นคืนข้อมูลจากชุดข้อมูลโอเพนดาทาแบบเปิดเผยรหัส โดยการเปรียบเทียบประสิทธิภาพจะใช้การวัดความเร็วในการประมวลผลจากชุดข้อมูลหลายขนาด ผลการทดสอบพบว่า โมเดลที่พัฒนาขึ้นโดยใช้การเตรียมข้อมูลล่วงหน้าและเทคนิคการค้นคืนเชิงความหมายพร้อมการวัดระยะทางเชิงความหมาย จะช่วยลดเวลาการประมวลผลคำค้นและค้นคืนรายการข้อมูลได้ในเวลาอันสั้นแม้ข้อมูลจะมีปริมาณมากก็ตาม แต่เจนาเฟรมเวิร์คจะใช้เวลาในการทำงานแปลผันตรงกับปริมาณข้อมูลโดยหากข้อมูลมีปริมาณมากก็จะใช้เวลามากขึ้นตามไปด้วย นอกจากนี้ เจนาเฟรมเวิร์คยังไม่สามารถรองรับการค้นคืนเชิงความหมายทำให้ประสิทธิภาพการค้นคืนต่ำกว่าโมเดลที่พัฒนาขึ้นจากการวิจัยนี้

2) การประเมินการใช้งานระบบที่พัฒนาขึ้นตามโมเดลโดยผู้ใช้ ใช้กลุ่มตัวอย่างทั้งหมด 20 คนในการทดสอบประสิทธิภาพการทำงานของระบบ ผลการประเมินพบว่า ระบบโดยรวมมีประสิทธิภาพในระดับดีมาก ($\overline{X} = 4.54$) และเมื่อพิจารณาแยกตามองค์ประกอบย่อยทั้ง 4 ด้าน พบว่า องค์ประกอบที่ได้รับการประเมินในระดับดีมาก ได้แก่ มอดูลการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์ และความง่ายในการใช้งานระบบโดยรวม ซึ่งมีค่าเฉลี่ยที่ 4.87 และ 4.54 ตามลำดับ ส่วนองค์ประกอบที่ได้รับการประเมินในระดับดี ได้แก่ มอดูลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ โดยทั้งสองส่วนมีค่าเฉลี่ยที่ 4.38 และ 4.37 ตามลำดับ

6.2 การประยุกต์ใช้ผลการวิจัย

6.2.1 การประยุกต์ใช้ผลวิจัยทางตรง คือการนำไปใช้เป็นกรอบการพัฒนาระบบเพื่อจัดเก็บและค้นคืนข้อมูลเลิร์นนิงอ็อบเจกต์ในวิชาคณิตศาสตร์ระดับชั้นมัธยมศึกษา โดยโมเดลการจัดเก็บสามารถรองรับการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์ที่อยู่ในระบบฐานข้อมูลที่มีชุดเมทาดาทาที่แตกต่างกัน เหมาะสำหรับผู้ใช้ที่มีข้อมูลเลิร์นนิงอ็อบเจกต์อยู่หลายแห่ง และใช้ชุดเมทาดาทาที่แตกต่างกันหลายมาตรฐาน ซึ่งในโมเดลใช้แนวคิดการผสานเมทาดาทา (Metadata Mapping) เพื่อผสานองค์ประกอบและโครงสร้างความสัมพันธ์ของตารางข้อมูลที่แตกต่างกันให้มีรูปแบบมาตรฐานกลาง ทั้งนี้แนวคิดในการผสานเมทาดาทาในงานวิจัยนี้ได้ใช้เทคนิคการผสานเมทาดาทา 2 ระดับได้แก่ ระดับโครงสร้างข้อมูล เพื่อจัดลำดับของคลาสที่เหมาะสม และการผสานระดับองค์ประกอบ (Element) โดยใช้ศัพท์สัมพันธ์เวิร์ดเน็ตในการเปรียบเทียบความคล้ายคลึงร่วมกับการวัดความคล้ายคลึงของแจ็คการ์ดในการหาค่าสัมประสิทธิ์ความคล้ายคลึง เพื่อใช้ในการพิจารณาการกำหนดคำศัพท์ของคลาสและคุณสมบัติ ทำให้การผสานเมทาดาทามีความถูกต้องมากขึ้น

ส่วนโมเดลการค้นคืนเลิร์นนิงอ็อบเจกต์ได้ใช้แนวคิดการสืบค้นเชิงความหมายโดยใช้วิธีการวัดความคล้ายคลึงของคำค้นและคำที่เกี่ยวข้องด้วยแนวคิดการสร้างเป็นตารางข้อมูลแบบคีย์-เวลู (Key-Value) และการจัดลำดับผลการสืบค้นด้วยการให้น้ำหนักองค์ประกอบที่ต้องการ ซึ่งสามารถประยุกต์แนวคิดการค้นคืนสามารถนำไปใช้ในการค้นคืนเชิงความหมาย เพื่อได้ผลการค้นที่มีรายการที่เกี่ยวข้องกับคำค้นและรายการที่เกี่ยวข้องตามความสัมพันธ์ของเนื้อหา ทำให้ได้ผลการค้นที่มีความถูกต้องและครอบคลุมเรื่องที่เกี่ยวข้องกับคำค้นได้แม่นยำมากกว่าการค้นด้วยการเปรียบเทียบเฉพาะคำค้นเพียงอย่างเดียว

ทั้งนี้การจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์ในโมเดลนี้ ออกแบบให้ใช้การจัดเก็บเป็นโครงสร้างแบบอาร์ดีเอฟ ซึ่งข้อด้อยของโครงสร้างนี้เมื่อมีข้อมูลปริมาณมากขึ้นจะทำให้การเข้าถึงข้อมูลทั้งหมดใช้เวลานานและส่งผลต่อประสิทธิภาพการค้นคืน ดังนั้นเทคนิควิธีการการค้นคืนจึงใช้การแปลงเป็นตารางข้อมูลเพื่อความรวดเร็ว ซึ่งใช้เทคนิคคีย์-เวลูจะแบ่งข้อมูลแต่ละส่วนของอาร์ดีเอฟออกเป็นคอลเล็กชั่นย่อย ๆ เพื่อสามารถเข้าถึงข้อมูลได้พร้อมกันทำให้ใช้เวลาค้นคืนน้อยกว่าการเข้าถึงข้อมูลอาร์อีเอฟทั้งชุดข้อมูลทีละระเบียน

6.2.2 การประยุกต์ใช้ผลวิจัยทางอ้อม โดยงานวิจัยได้นำแนวคิดการผสานเมทาดาทาโดยใช้คำศัพท์สัมพันธ์มาใช้ในการวิเคราะห์ความคล้ายคลึงร่วมกับการวัดค่าสัมประสิทธิ์ความคล้ายคลึงจากทฤษฎีนักวิจัยหลายท่าน โดยใช้วิธีการวัดที่สอดคล้องโครงสร้างข้อมูลที่จัดเก็บ รวมถึงการผสานเมทาดาทาในงานวิจัยนี้จะทำ 2 ระดับ ซึ่งสามารถออกแบบกระบวนการให้ทำการพิจารณาข้อมูลในองค์ประกอบเพิ่มขึ้น

ได้ เช่น การพิจารณาชนิดข้อมูล (Data Type) และค่า (Value) ของข้อมูล เป็นต้น ทำให้การผสานเมทาดาทามีความแม่นยำและถูกต้องมากยิ่งขึ้น

นอกจากนี้แนวคิดการค้นคืนในชุดข้อมูลโอเพนดาทาผู้วิจัยได้ใช้การวัดความคล้ายคลึงโดยใช้เทคนิคการเปรียบเทียบคำค้นกับศัพท์สัมพันธ์ก่อนทำการค้นคืนร่วมกับการกำหนดขอบเขตของหัวเรื่องที่เกี่ยวข้องในการค้นคืนจะช่วยเพิ่มประสิทธิภาพในการค้น ทำให้การค้นคืนในโครงสร้างอาร์ดีเอฟมีประสิทธิภาพมากกว่าการค้นแบบเปรียบเทียบโดยตรงจากการค้นแบบปกติ ซึ่งเป็นแนวคิดที่ช่วยเพิ่มประสิทธิภาพการค้นคืนเชิงความหมายในโครงสร้างอาร์ดีเอฟอีกรูปแบบหนึ่งที่เป็นทางเลือกในการเข้าถึงข้อมูลอาร์ดีเอฟที่มีข้อมูลปริมาณมากได้ดี

6.3 ข้อจำกัดและข้อเสนอแนะในการทำวิจัย

จากการศึกษาและพัฒนาโมเดลการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์จากชุดข้อมูลในลิงค์โอเพนดาทามีข้อจำกัดในการทดลองและเสนอแนะ ดังต่อไปนี้

6.3.1 ข้อจำกัดในการวิจัย

1) งานวิจัยนี้มีการควบคุมขนาดชุดข้อมูล ดังนั้น การนำชุดข้อมูลที่ใช้งานจริงที่มีการเชื่อมโยง URI จำนวนมากมาทดสอบ ต้องมีการศึกษาเพิ่มเติมถึงขีดจำกัดในการประมวลผลของเครื่องมือทดลอง และประสิทธิภาพของเทคนิคการวัดความคล้ายคลึงอื่น ๆ เพิ่มเติม เพื่อป้องกันไม่ให้เกิดความเสียหายกับระบบทดสอบ

2) แนวคิดการวัดความคล้ายคลึงเชิงความหมาย ประกอบด้วยการวัดระยะทางเชิงความหมายและความคล้ายคลึงเชิงความหมายของคำศัพท์ ออกแบบเพื่อใช้กับชุดข้อมูลในงานวิจัยนี้เท่านั้น การนำไปใช้กับชุดข้อมูลอื่น ต้องมีการศึกษาแนวคิดอื่น ๆ เพิ่มเติม และอาจมีการออกแบบกระบวนการเปรียบเทียบประสิทธิภาพในการค้นคืนเชิงความหมายแนวคิดอื่น ๆ เพื่อหาแนวคิดที่เหมาะสมกับประเภทและลักษณะโครงสร้างข้อมูลในแต่ละชุดข้อมูล ก่อนจะเลือกใช้งาน

3) แนวทางในการผสานเมทาดาทามีความจำเพาะกับเค้าร่างของเมทาดาทาแต่ละมาตรฐาน ดังนั้น แนวทางในอนาคตอาจมีกระบวนการในการปรับเปลี่ยนชุดคำศัพท์ หรือรองรับชุดเค้าร่างเมทาดาทาที่มีโดเมนอื่น ๆ ภายในระบบเดียวกันได้ โดยใช้แนวคิดการผนวกออนโทโลยี เป็นต้น

4) โมเดลการจัดเก็บเลิร์นนิงอ็อบเจกต์รองรับการนำเข้าข้อมูลในรูปแบบฐานข้อมูลเชิงสัมพันธ์และไฟล์เอสคิวแอล (SQL) เท่านั้น

5) ข้อมูลที่ใช้ในการทดสอบระบบจำเป็นจะต้องกรอกข้อมูลให้ครบถ้วนตามที่ระบบกำหนด เพื่อให้การทำงานของระบบมีประสิทธิภาพตรงตามที่ออกแบบไว้

6.3.2 ข้อเสนอแนะในการวิจัยครั้งต่อไป

1) แนวคิดการผสานเมทาดาทา ควรออกแบบให้รองรับการผสานเมทาดาทาในมาตรฐานอื่น ๆ ได้ หรือรองรับกลุ่มคำศัพท์ในโดเมนที่แตกต่างกัน หรือมีขอบเขตโดเมนที่กว้างขึ้น เช่น กลุ่มสาระวิชาอื่น หรือระดับชั้นเรียนนอกเหนือจากระดับมัธยมศึกษา เป็นต้น

2) การวัดความคล้ายคลึงเชิงความหมาย สามารถปรับปรุงอัลกอริทึมในการกำหนดขอบเขตของรายการที่เกี่ยวข้องด้วยเทคนิคอื่น ๆ นอกเหนือจากการพิจารณาเมทาดาทา เช่น การพิจารณาถึงข้อมูลหรือชนิดข้อมูลประกอบด้วยจะช่วยให้การค้นคืนมีประสิทธิภาพยิ่งขึ้น

3) การเข้าถึงไฟล์อาร์ดีเอฟที่มีปริมาณข้อมูลมากจะใช้เวลามากขึ้น ดังนั้นเทคนิควิธีการเพิ่มความรวดเร็วในการเข้าถึงข้อมูลในไฟล์อาร์ดีเอฟยังสามารถพัฒนาได้อีกหลายวิธี

มหาวิทยาลัยเทคโนโลยีสุรนารี

รายการอ้างอิง

กรมวิชาการ กระทรวงศึกษาธิการ. (2545). **เอกสารประกอบหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 การวิจัยเพื่อพัฒนาการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐาน**. กรุงเทพฯ: โรงพิมพ์คุรุสภาลาดพร้าว.

กรมวิชาการ กระทรวงศึกษาธิการ. (2555). **นโยบายกระทรวงศึกษาธิการ**. [ออนไลน์]. ได้จาก: www.moe.go.th/moe/th/news/detail.php?NewsID=26731&Key=news20

กิดานันท์ มลิทอง. (2548). **เทคโนโลยีการศึกษาและนวัตกรรม**. พิมพ์ครั้งที่ 2. กรุงเทพฯ: อรุณการพิมพ์.

โครงการมหาวิทยาลัยไซเบอร์ไทย. (2556). **ระบบสืบค้นทรัพยากรการเรียนรู้นานาชาติ.** [ออนไลน์]. ได้จาก: http://globe.thaicyberu.go.th/

ใจทิพย์ ณ สงขลา. (2550). **E- Instructional Design วิธีการออกแบบการเรียนการสอนอิเล็กทรอนิกส์.** กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ชัชวาลย์ วงษ์ประเสริฐ. (2537). **บริการสารนิเทศ**. กรุงเทพฯ: สาขาสารนิเทศ คณะนิเทศศาสตร์ มหาวิทยาลัยรังสิต.

ชาญชัย พิพัฒน์สันติกุล. (2548). สภาพเทคโนโลยีการศึกษาไทยในปี พ.ศ. 2550 ตามความคาดการณ์ของนักเทคโนโลยีการศึกษา. วิทยานิพนธ์ครุศาสตรมหาบัณฑิตจุฬาลงกรณ์มหาวิทยาลัย.

ถนอมพร เลาหจรัสแสง. (2550). **นิยามเลิร์นนิง อ็อบเจกต์ เพื่อการออกแบบพัฒนาสื่ออิเล็กทรอนิกส์.** วารสารเทคโนโลยีและสื่อสารการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช. 4(4): 50-59.

นฤมล ปราชญโยธิน. (2545). **การประเมินระบบจัดเก็บและค้นคืนสารสนเทศ ใน ประมวลสาระชุดวิชาการจัดเก็บและการค้นคืนสารสนเทศ หน่วยที่ 7-8**. นนทบุรี: สาขาวิชาศิลปศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช.

นฤมล ปราชญโยธิน ทวีศักดิ์ กออนันตกูล และ เปรมิน จินดาวิมลเลิศ. (2536). **ธิซอรัสกับระบบสารสนเทศ.**กรุงเทพฯ: ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ.

นฤมล ปราชญโยธิน. (2556). **คลังคำในบริบทการจัดเก็บและค้นคืนสารสนเทศ**. มหาสารคาม: หจก.อภิชาติการพิมพ์.

นฤมล รักษาสุข. (2548). **เอกสารประกอบการเรียนวิชา 204213 การจัดการทรัพยากรสารสนเทศ**. นครราชสีมา : สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี.

นิศาชล จำนงศรี. (2554). บทวิจารณ์หนังสือ Metadata and Its Applications in the Digital Library: Approaches and Practices (By Jia Liu). **วารสารเทคโนโลยีสุรนารี.** 5 (1): 77-81.

นิศาชล จำนงศรี. (2546). **เทคโนโลยีในงานสารสนเทศ**. ฉบับปรับปรุงครั้งที่ 3. นครราชสีมา: สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี.

บุญชม ศรีสะอาด. (2553). **การวิจัยเบื้องต้น**. กรุงเทพฯ: สุวีริยาสาร์น.

บุญดี บุญญากิจ, นงลักษณ์ ประสพสุขโชคชัย, ดิสพงศ์ พรชนกนาถ และปรียวรรณ กรรณล้วน. (2549). **การจัดการความรู้จากทฤษฎีสู่การปฏิบัติ (พิมพ์ครั้งที่ 3)**. กรุงเทพฯ: จิรวัฒน์ เอ็กเพรส.

บุญเลิศ อรุณพิพูลย์. (2547). E-learning ในประเทศไทย. **เนคเทค**. 11 (56): 32-26.

ประดิษฐา ศิริพันธ์. (2543). กลยุทธ์การสืบค้นสารสนเทศ. **รังสิตสารสนเทศ มกราคม-มิถุนายน**. 6 (1): 25-32.

ประดิษฐา ศิริพันธ์. (2546). **การประยุกต์ดับลินคอร์เมทาดาทาเพื่อการสืบค้นสารสนเทศ**. [ออนไลน์]. ได้จาก: http://www.tiac.or.th/duplincore_apply2003.pdf

ถนอมพร เลาหจรัสแสง. (2550). **นิยามเลิร์นนิง อ็อบเจกต์ เพื่อการออกแบบพัฒนาสื่ออิเล็กทรอนิกส์.** วารสารเทคโนโลยีและสื่อสารการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช. 4 (4): 50-59.

พระบาทสมเด็จพระเจ้าอยู่หัวภูมิพลอดุลยเดช. (2543). **พระราชธรรม พระราชประวัติ พระราชกรณียกิจ พระบาทสมเด็จพระเจ้าอยู่หัวภูมิพลอดุลยเดชมหาราช**. กรุงเทพฯ: ธรรมสภา.

เพ็ญพรรณ อัศวนพเกียรติ, อรินทิพย์ ธรรมชัยพิเนต และ กฤษณะ ไวยมัย. (2546). ออนโทโลยีชีวภาพ: ระบบสาหรับสืบค้นและวิเคราะห์ข้อมูลทางด้านชีววิทยา. **เอกสารประกอบการประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 41: สาขาพืช สาขาวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์** (หน้า 277-285). กรุงเทพ: มหาวิทยาลัยเกษตรศาสตร์.

มนต์ชัย เทียนทอง. (2548). **การพัฒนาบทเรียนคอมพิวเตอร์ช่วยสอนระบบมัลติมีเดียสำหรับฝึกอบรมครู-อาจารย์และนักฝึกอบรม เรื่อง การสร้างบทเรียนคอมพิวเตอร์ช่วยสอน**. วิทยานิพนธ์ปริญญาครุศาสตร์อุตสาหกรรมดุษฎีบัณฑิต. สาขาวิจัยและพัฒนาหลักสูตร

ภาควิชาบริหารเทคนิคศึกษา บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

มารุต บูรณรัช และ เทพชัย ทรัพย์นิธิ. (2553). **การอบรมเชิงปฏิบัติการ: พัฒนาออนโทโลยีเพื่อสนับสนุนห่วงโซ่อุปทานกล้วยไม้.** กรุงเทพฯ: ศูนย์เทคโนโลยีอิเล็คทรอนิกส์และคอมพิวเตอร์แห่งชาติ.

มาลี กาบมาลา ลำปาง แม่นมาตย์ และ ครรชิต มาลัยวงศ์. (2549). ออนโทโลยี: แนวคิดการพัฒนา. **บรรณรักษศาสตร์และสารนิเทศศาสตร์ มข.** 24 (1-3): 24-49.

มาลี กาบมาลา. (2553). **เอกสารประกอบการสอนวิชา 412 231 บริการสารสนเทศ.** ขอนแก่น: สาขาการจัดการสารสนเทศและการสื่อสาร คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น.

มาลี ล้าสกุล. (2549). **การบริการและเผยแพร่สารนิเทศ.** นนทบุรี: แขนงวิชาสารสนเทศศาสตร์ สาขาวิชาศิลปศาสตร์ มหาวิทยาลัยสุโขทัยธรรมาธิราช.

รวีวรรณ ขาพล และ นริศรา เฮมเบีย. (2551). การบริการตอบคาถามและช่วยการค้นคว้า: ความคาดหวังและความพึงพอใจของผู้ใช้บริการ ฝ่ายหอสมุดจอห์น เอฟ เคนเนดี้. **วารสารวิชาการคณะมนุษยศาสตร์และสังคมศาสตร์.** 4 (1): 104-132.

รัญจวน อินทรกำแหง. (2517). **วิชาครูประกาศนียบัตรครูมัธยมของคุรุสภาวิชาบรรณารักษศาสตร์.** กรุงเทพฯ: คุรุสภา

โรสริน อัคนิจ ธนิต พุทธพงศ์ศิริพร น้าฝน ลาดับวงศ์ และ อารีย์ ธัญกิจจานุกิจ. (2554). การพัฒนาออนโทโลยีเพื่อการจัดการความรู้ด้านการแปรรูปข้าว. **วารสารเกษตร.** 27 (3): 267-274.

วรวิทย์ พงษ์เกษมวิวัฒน์. (2551). ระบบสืบค้นเอกสารงานวิจัยเชิงความหมาย. **วารสารวิชาการพระจอมเกล้าพระนครเหนือ.** ฉบับที่ 18

วรสุดา บุญไวโรจน์. (2540). **การพัฒนาทักษะทางคณิตศาสตร์ในระดับประถมศึกษา ในเรื่องน่ารู้สำหรับครูคณิตศาสตร์.** กรุงเทพ ฯ: ไทยวัฒนาพานิช.

วิชชุ พร้อมจรรยากุล. (2547). **การออกแบบและพัฒนาโมบายเอเจนต์ด้วยภาษาจาวา.** กรุงเทพฯ: บัณฑิตวิทยาลัยเกษตรศาสตร์.

วิชุดา โชติรัตน์ ผุสดี บุญรอด และ ศจีมาจ ณ วิเชียร. (2554). การพัฒนาฐานความรู้ออนโทโลยีสำหรับวิเคราะห์ข่าวออนไลน์โดยอัตโนมัติ. **วารสารเทคโนโลยีสารสนเทศ.** 7 (17): 13-18.

วิสิทธิ์ บุญชุม และ นวลวรรณ สุนทรภิษัช. (2552). ออนโทโลยี: สื่อกลางในการจับคู่ข้อมูลเอ็กซ์เอ็มเอลที่มีความสัมพันธ์กันเชิงความหมาย. **เอกสารประกอบการสัมนาทางวิชาการ The**

5th National Conference on Computing and Information Technology-NCCIT 2009. (หน้า 338-343). กรุงเทพฯ: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

วัฒนา เฉียงเหนือ. (2549). **สสวท. เปิดตัวนวัตกรรมใหม่ Learning Object สื่อดิจิตัลศักยภาพสูง เพื่อการเรียนรู้ รุกคืบสู่ชั้นเรียนวิทย์-คณิต ทั่วประเทศไม่นานเกินรอ.** [ออนไลน์]. ได้จาก: http://www.ipst.ac.th

ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. (2555). **เครือข่ายกาญจนาภิเษก**. [ออนไลน์]. ได้จาก: http://kanchanapisek.or.th/speeches/index.th.html

ศยามน อินสะอาด. (2550). **การออกแบบและพัฒนาอีเลิร์นนิ่ง**. พิมพ์ครั้งที่ 1. นครราชสีมา: มหาวิทยาลัยเทคโนโลยีสุรนารี.

สติยา ลังการ์พินธุ์. (2548). Learning Object: สื่อการเรียนรู้ยุคดิจิตอล. **สสวท**. 33 (1-2): 70-74.

สถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.). (2555). **ค่าสถิติพื้นฐาน O-NET รายวิชา จำแนกตามเขตพื้นที่การศึกษา ป.6 ปีการศึกษา 2550-2553**. [ออนไลน์]. ได้จาก: http://www.niets.or.th/index.php/research_th

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2549). **สสวท. พัฒนาสื่อดิจิตัล Learning Object**. [ออนไลน์] ได้จาก: http://www.ipst.ac.th

สมชาย ปราการเจริญ. (2548). ออนโทโลยีทางเลือกของการพัฒนาฐานความรู้ในรูปแบบเชิงเนื้อหา. **เอกสารประกอบการสัมมนาทางวิชาการ The 5th National Conference on Computing and Information Technology-NCCIT 2009.** (หน้า 92-99). กรุงเทพฯ: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

สมพงษ์ ปั้นหุ่น. (2552). **สรุปผลการประเมินคุณภาพการศึกษาระดับนานาชาติและปัจจัยที่เกี่ยวข้องเพื่อเสนอแนะแนวทางการยกระดับคุณภาพการศึกษาด้านวิทยาศาสตร์ คณิตศาสตร์และการอ่าน**. กรุงเทพฯ: สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน.

สมมณี ลูซะวงษ์ และ งามนิจ อาจอินทร์. (2553). การเข้าถึงฐานข้อมูลบนพื้นฐานของออนโทโลยีด้วยวิธีการปรับเปลี่ยนคาสั่งการสืบค้น. **เอกสารประกอบการสัมมนาทางวิชาการ The Second Conference on Knowledge and Smart Technologies** (หน้า 29-34). ชลบุรี: ประเทศไทย.

สมฤดี วิศทเวทย์. (2536). **ทฤษฎีความรู้ของฮิวม์**. กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สมวงษ์ แปลงประสพโชค, สมเดช บุญประจักษ์ และ จรรยา ภูอุดม. (2549). **นวัตกรรมเพื่อพัฒนาคุณภาพการจัดการศึกษาคณิตศาสตร์ขั้นพื้นฐานของเด็กไทย: การศึกษาสาเหตุเด็กไทยอ่อนคณิตศาสตร์และแนวทางแก้ไข.** มหาวิทยาลัยราชภัฏพระนคร.

สมวงษ์ แปลงประสพโชค. (2549). ปัญหาการเรียนการสอนคณิตศาสตร์และแนวทางแก้ไข. **วารสารวงการครู**. 3 (31): 78-80.

สมาคมห้องสมุดแห่งประเทศไทย. (2525). หัวเรื่องสำหรับหนังสือภาษาไทย. พิมพ์ครั้งที่ 4 แก้ไขปรับปรุง. กรุงเพทมหานคร: สมาคมห้องสมุดแห่งประเทศไทย.

สำนักงบประมาณ. (2554). **งบประมาณโดยสังเขป ประจำปีงบประมาณพ.ศ.2554**. กรุงเทพฯ: รุ่งศิลปการพิมพ์.

สำนักงานข้าราชการครูและบุคลากรทางการศึกษา. (2551). **ระบบฐานข้อมูลข้าราชการครูและบุคลากรทางการศึกษา**. [ออนไลน์]. ได้จาก: http://203.146.15.34/moe_teacher/report_gov.php

สำนักงานคณะกรรมการการศึกษาแห่งชาติ. (2543). **ยุทธศาสตร์การเรียนรู้ตลอดชีวิตในศตวรรษที่ 21**. กรุงเทพฯ: สกศ.

สำนักงานคณะกรรมการการศึกษาแห่งชาติ. (2554). **แผนการศึกษาแห่งชาติ (พ.ศ. 2552-2559)**. กรุงเทพ: พริกหวานกราฟิก.

สำนักงานคณะกรรมการการศึกษาเอกชน. (2542). **การจัดการเรียนการสอนโดยใช้ภาษาอังกฤษเป็นสื่อการสอนในโรงเรียนเอกชน**. กรุงเทพฯ: สำนักงานคณะกรรมการการศึกษาเอกชน.

สำนักงานคณะกรรมการการส่งเสริมการศึกษาเอกชน. (2555). **โรงเรียนเอกชนที่จัดการเรียนการสอนตามหลักสูตรกระทรวงศึกษาธิการเป็นภาษาอังกฤษ (English Program)**. [ออนไลน์]. ได้จาก: http://www.opec.go.th/index.php?option=com_content&view=article&id=4546%3A-english-program&catid=14%3A2009-10-04-12-26-14&Itemid=12&lang=en

สำนักงานคณะกรรมการพัฒนาระบบราชการ. (2546). **คู่มือคำอธิบายและแนวทางปฏิบัติงานพระราชกฤษฎีกาว่าด้วยหลักเกณฑ์และวิธีการบริหารกิจการบ้านเมืองที่ดี**. กรุงเทพมหานคร: สิริบุตรการพิมพ์.

สำนักงานเลขาธิการสภาการศึกษา. (2553). **แผนการศึกษาแห่งชาติ ฉบับปรับปรุง (พ.ศ.2552-2559)**. กรุงเทพฯ: สกศ.

สำนักงานสถิติแห่งชาติ. (2551). **สำรวจการมีการใช้เทคโนโลยีสารสนเทศและการสื่อสารในสถานศึกษาพ.ศ. 2551**. กรุงเทพฯ: บางกอกบล๊อก.

สิริรัตน์ ประกฤติกรชัย. (2550). **การสร้างต้นแบบออนโทโลยีของพืชสมุนไพรไทย.** วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

สีลตา วงศ์กาฬสินธุ์ และ งามนิจ อาจอินทร์. (2552). การรวมกันของออนโทโลยีอย่างมีความหมายสาหรับโดเมนห้องสมุดอิเล็กทรอนิกส์. **เอกสารประกอบการสัมนาทางวิชาการ The 5th National Conference on Computing and Information Technology-NCCIT 2009** (หน้า 288-293). กรุงเทพฯ: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

สุทธิลักษณ์ อำพันวงศ์. (2510). **หลักเกณฑ์การทำบัตรรายการหนังสือภาษาไทยฉบับสมบูรณ์และตัวอย่างบัตร**. พระนคร: ไทยวัฒนาพานิช

สุวิมล สุทธนะ. (2545)**.** การสร้างบทเรียน e-learning เพื่อให้เกิดผลสัมฤทธิ์ทางการศึกษา. **วารสารรามคำแหง**. 4 (19): 19-23.

สุมิตรา จิระวฒุินันท์. (2543). **แหล่งข้อมูลทางด้านวิศวกรรมศาสตร์จากอินเตอร์เน็ต**. ขอนแก่น: มหาวิทยาลัยขอนแก่น.

อนุชัย ธีระเรืองชัยศรี. (2549). **Learning Object.** [ออนไลน์]. ได้จาก: http://www.learningsquare .com/download/seminar2007/Anuchai_present_3.ppt

อนุชัย ธีระเรืองชัยศรี. (2551). **Learning Object: จากมุมมองด้านการเรียนการสอน**. [ออนไลน์]. ได้จาก: http://www.learnsquare.com/download/seminar2007/Anuchai_ present_3. ppt

อรวรรณ อุไรเรืองพันธุ์ และ สมจิตร อาจอินทร์. (2552). การสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยี. **เอกสารประกอบการสัมนาทางวิชาการ The 5th National Conference on Computing and Information Technology-NCCIT 2009** (หน้า 294-299). กรุงเทพฯ: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

อิศราวดี ทองอินทร์. (2553). การพัฒนาบริการสารสนเทศออนไลน์เพื่อการวิจัยในสาขาวิชาพระพุทธศาสนา มหาวิทยาลัยมหาจุฬาลงกรณ์ราชวิทยาลัย วิทยาเขตเชียงใหม่. **สารสนเทศศาสตร์.** 28 (1): 1-15.

อุทิตย์ พิมพา. (2545). การพัฒนาสารสนเทศด้านกฎหมายบนเครือข่ายคอมพิวเตอร์ Metadata Dublin Core ศูนย์สนเทศและห้องสมุด มหาวิทยาลัยธุรกิจบัณฑิตย์. **บรรณารักษศาสตร์**. 8 (1): 23-38.

Alahmari, F., Thom, J., Magee, L. and Wong, W. (2012). Evaluating semantic browsers for consuming linked data. In **Proceedings of the Twenty-Third Australasian Database Conference** (pp 89–98). Melbourne, Australian Computer Society.

AnHai, D. Domingos, P. and Halevy, A. (2003). Learning to Match the Schemas of Data Sources: A Multistrategy Approach. in **Machine Learning**. 50 (3): 279-301.

AnHai, D. Jayant, M., Domingos, P.; and Halevy A. (2003). Learning to Map between Ontologies on the Semantic Web. **VLDB Journal, Special Issue on the Semantic Web** (pp 626-673). New York: Association Computer Machinery.

Bae, M., Nguyen, D., Kang, S. and Oh, S. (2013). K-depth RDF keyword search algorithm based on structure indexing. In **Advance Methods Technology Agent Multi-Agent System** (pp 346-355). Hue City: IOS Press.

Baeza-Yates, R. and Ribeiro-Neto, B. (2005). **Modern Information Retrieval**. New York: Addison-Wesley.

Baeza-Yates, R. and Ribeiro-Neto, B. (2011). **Modern Information Retrieval 2nd edition**. New York: Addison-Wesley.

Balatsoukas, P., Morris, A., and O'Brien, A. (2008). Learning Objects Update: Review and Critical Approach to Content Aggregation. **Educational Technology & Society**. 11 (2): 119-130.

Barcelos, C. F. and Gluz, J. C. (2011). An agent based federated learning object search service. **Interdisciplinary Journal of E-Learning and Learning Objects**. 7 (1): 37-54

Berners-Lee, T. (2000). **Semantic Web on XML**. [On-line]. Available: http://www.w3.org/2000 /Talks/1206-xml2k-tbl/slide10-0.html

Berners-Lee, T. (2006). **Linked Data –Design Issue**. [On-line]. Available: http://www.w3c.org/ Designissue/linkedData.html.

Berners-Lee, T., Christian, B. and Heath, T. (2009). Linked data the story so far. **International Journal on semantic web and information system**. 5 (3): 1-22.

Bizer, C. (2009). The emerging web of linked data. **Intelligent Systems**. IEEE. 25 (5): 87–92.

Blake, C. and Pratt, W. (2001). Better Rules, Fewer Features: A Semantic Approach to Selecting Features from Text. In **conjunction with the First IEEE Data Mining Conference, Workshop on Text Mining** (pp 59-66). New York: Institute of Electrical and Electronics Engineers.

Bogdan A., Mauricio H., Lucian P. and Wang-Chiew T. (2010). MapMerge: Corerelating independence schema mapping. In **Proceeding of the 36th International Conference on Very Large Data Bases** (pp 81-92). Singapore: Very large database endowment.

Bollacker, K. D., Lawrence, S. and Giles, C. L. (2000). Discovery Relevant Scientific Literature on The Web. In **IEEE Intelligent Systems**. 15 (20): 42 - 47.

Bouquet, P., Giunchiglia, F., Harmelen, F. V., Serafini, L. and Stuckenschmidt. H. (2003). C-OWL: Contextualizing Ontologies. In **Proceeding of 2nd International Semantic Web Conference 2003** (pp 164-179). New York: Springer Berlin Heidelberg.

Bouzeghoub, A. and Elbyed, A. (2006). Ontology mapping for learning objects repositories interoperability. In **Proceedings of 8th Intelligent Tutoring Systems. 2006** (pp 794-797). New York: Springer Berlin Heidelberg.

Bradshaw, J. M. (1997). **Software agent**. California: American Association for Artificial Intelligence.

Brank, J., Grobelnik, M. and Mladenic, D. (2005). **A Survey of Ontology Evaluation Techniques. Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD).** Ljubljana: Slovenia.

Brickley, D. and Guha, R.V. (2008). **RDF Vocabulary Description Language 1.0: RDF Schema**. [On-line]. Available: http://www.w3.org/TR/rdf-schema/

Brin, S. and Page, L. (1998). The anatomy of a large-scale hyper textual web search engine. **Computer Networks and ISDN Systems**. 30 (1-7): 107-117.

Broughton, V. (2006). **Essential thesaurus construction.** London: Facet Publishing.

Burkowski, F. (1992). Retrieval activities in a database consisting of heterogeneous collections of structured texts. In **Proceedings of the 15th ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'92)** (pp 112-125). New York, Association for Computing Machinery Press Press.

Cai, D., Yu, S., Wen, R. and Ma, Y. (2003). **VIPS: A vision-based page segmentation algorithm**. [On-line]. Available: https://www.microsoft.com/en-us/research/publication/vips-a-vision-based-page-segmentation-algorithm/

Candan K. S., Kim J. W., Liu H. and Suvarna R. (2006). Discovering mappings in hierarchical data from multiple sources using the inherent structure. **Knowledge and Information System**. 10 (2): 185–210.

Caplan, P. (2003). **Metadata fundamentals for all librarians**. Chicago: American Library Association.

Changqing, L., Ling, T. W. (2004). OWL-based semantic conflicts detection and resolution for data interoperability. In **Proceeding of ER workshop 2004.** (pp 266-277). New York: Springer Berlin Heidelberg

Chen, J., Zhou, B., Shi, J., Zhang, J. and Qiu, F. (2001). Function-based object model towards website adaptation. In **Proceedings of the 10th International World Wide Web Conference** (pp 587-596). New York: Association for Computing Machinery Press Press.

Choo, C. W. (2000). Working knowledge: how organization manage what they know. **Library management**. 21 (8): 395-403.

Christain, B. and Radoslaw, O. (2004). **RAP: RDF API for PHP**. Computer and Information Science Papers.

Citec CU. (2011). **BCogS**. [On-line]. Available: http://www.cit-ec.de/sites/www.cite-ec.de/files /slides_cunger.pdf

Clarke, C., Cormack G. and Burkowski, F. (1995). Algebra for structured text search and a framework for its implantation. **The Computer Journal**. 38 (1): 43-56.

Claudia, L. and Chodorow, M. (1998). **WordNet An Electronic Lexical Database.** Cambridge: MIT press.

Cronbach, L. J. (1970). **Essentials of Psychological Test**. (5th ed.). New York: Harper Collins.

Cullot, N., Ghawi, R. and Yetongnon, K. (2007). DB2OWL: a tool for automatic database-to-ontology mapping. **Proceedings of 15th Italian Symposium on Advanced Database System (SEBD 2007)** (pp 491-494). New York: Springer Berlin Heidelberg.

Currás, E. (2010). **Ontologies, taxonomies and thesauri in systems science and systematic.** Chandos: Oxford University Press.

Cutting, D. R., Karger, D. R., Pedersen, J. O. and Tukey, J. W. (1992). Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections. In **Proceeding of the 15th annual international ACM SIGIR conference on Research and development in information retrieval**. (pp 318-329). New York: Association for Computing Machinery Press.

Cyganiak, R. (2014). **The linked Open Data Cloud Diagram**. [On-line]. Available: http://lod-cloud.net

D2RQ. (2012). **Getting start with D2RQ**. [On-line]. Available: http://d2rq.org/getting-started

Daft, R. L.; and Lengel, R. H. (1984). Information richness: a new approach to managerial behavior and organization design. **Research in Organizational behavior**. 6 (5): 191-233.

Darlington, K. (2000). **The essence of expert systems**. Prentice Hall: New York.

Davenport, T. and Prusak, L. (1998). **Working knowledge: how organization manage what they know.** Cambridge: Harvard Business School Press.

Dawn, G. G. (2007). E-learning agent. **The learning Organization**. 4 (14): 300-312.

DCMI, (2013), **Dublin Core Metadata Element Set, Version 1.1**. [On-line]. Available: http:// dublincore.org/documents/dces/

Delbru, R., Capinas, S. and Tummerello, G. (2012). Searching web data: an entity retrieval and high-performance indexing model. **Web Semantics: Science, Services and Agents on the World Wide Web**. 10 (1): 33–58.

Dietze, H. and Schroeder, M. (2009). GoWeb: a semantic search engine for the life science web. **BMC bioinformatics**. 10 (1): 7.

Dragan, G. Jovanovic, J. and Vladan, D. (2007). Ontology-based annotation of learning object content. **Interactive Learning Environment**. 15 (1): 1-26.

Drucker, P. 1999. **Management challenge for the 21st century**. New York: harper business.

DSpace. (2014). **Top Reasons to Use DSpace**. [On-line]. Available: http://www.dspace.org/why-use.

Ellis, D. and Haugan, M. (1997). Modelling the information seeking patterns of engineering research scientists. in **Industrial Environment Journal of Documentation**. 53 (4): 384-403.

Ermine, J. (1995). **Expert systems: theory and practice**. New Delhi: Prentice Hall.

Fallows, D. (2005). **Search engine users**. Washington, D.C.: Pew Internet & American Life Project.

Feather, J. and Sturges, P. (1997). Information-seeking behaviour. In **International encyclopedia of information and library science**. New York: TJ Press.

Fenichel, C. H. and Hogan, T. H. (1981). **Online searching: A Primer**. Marlton, New jersey: Learned Information.

Fugmann, R. (1993). **Subject analysis and indexing: theoretical foundation and practical advice**. Frankfurt/Main, Germany: Indeks Verlag.

มหาวิทยาลัยเทคโนโลยีสุรนารี

Genesereth, M.R. and Kechpel, S.P. (1994). Software Agents. **Communication of the ACM**. 37 (7): 48-53.

Gonçalves, B., Meiss, M. R., Ramasco, J. J., Flammini, A. and Menczer, F. (2009). Remembering what we like: Toward an agent-based model of web traffic. In **WSDM'09: Proceedings of the 2nd ACM International Conference on Web Search and Data Mining**.

Good, C.V. (1973). **Dictionary of education** (3rd ed). USA: McGraw-Hill.

Gragan G. and Hatala M. (2006). Ontology mapping to improve learning resource search. **British Journal of Educational Technology (BJET)**. 37 (3): 375-390.

Gruber, R. T. (1993). A Translation Approach to Portable Ontology Specifications. **Knowledge Acquisition,** 5 (2): 199-220.

Gruber, R. T. (1993). Toward principles for the design of ontologies used for knowledge sharing. Presented at the Padua workshop on Formal Ontology. **International Journal of Human-Computer Studies**. 43 (4): 907-928.

Guarino N. (1995). Formal Ontology, Conceptual Analysis and Knowledge Representation. **International Journal of Human-Computer Studies**. 43 (5-6): 625–640.

Hagedorn, S. (2013). Discovery querying in Linked Open Data. In **Proceeding of 16th International Conference on Extending Database Technology** (pp 38-44). New York: Association for Computing Machinery Press.

Halevy, A. Y. (2001). Answering queries using views: A survey. **Very Large Databases Journal**. 10 (4): 270–294.

Hamid, S. and Sattar, C. A. (2002). Using Dewey decimal classification scheme (DDC) for building taxonomies for knowledge organization. **Journal of Documentation**. 58 (5): 575-583.

Haslholfer, B. (2008). A comparative study of mapping solutions. [On-line]. Available: http://www .cs.univie.ac.at/publication.php?pid=3886

Hassan, U., Sean, O. and Curry, E. (2012). Leveraging matching dependencies for guided user feedback in Linked Data application. In **Proceeding of 9th International Workshop on Information Integration on The Web** (Article No.5). New York: Association for Computing Machinery Press.

Heeptaisong, T. and Srivihok, A. (2010). Ontology Development for Searching Soil Knowledge. In **Proceeding of The 9th International Conference on e-Business 2010** (pp 102-107). Bangkok: Faculty of Science, Kasetsart University.

Hendler, J. (2001). Agents on the Web. **IEEE Intelligent Systems**. 16 (2): 30-37.

Hernandes, N. (2008). A model to represent the facets of learning objects. **Journal of E-learning and learning Object**. 4: 65-82.

Herner, S. (1970). Browsing. In Allen Kent, Harold, Lancour, and William Z. Nasri. (Eds.), **Encycropedia of library and information science**. New York: Marcel Dekker.

Hillmann, D. (2001). **Dublin Core Metadata Initiative: Using Dublin core**. [On-line]. Available: http://dublincore.org/documents/usageguide/index.shtml

Hjorland, B. (1997). **Information seeking and subject representation**. London: Greenwood Press.

Hogan, A., Umbrich, J., Harth, A., Cyganiak, R., Polleres, A. and Decker, S. (2012). An empirical survey of Linked Open Data conformance. **Web Semanics: Science, Services, and Agents on the world wide web**. 14 (12): 14-44.

HP Labs Semantic Web Programme. (2008). **Jena – A Semantic Web Framework for Java**. [On-line]. Available: http://jena.sourceforge.net/

Hsu, I. C. (2012). Intelligent discovery for learning objects using semantic web technologies. **Educational Technology and Society**. 15 (1): 298-312.

IEEE Learning Technology Standards Committee (LTSC). (2003). **IEEE P1484.12 Learning Object Metadata Working Group.** [On-line]. Available: http://ltsc.ieee.org/wg12/index.html.

IEEE LTSC. (2002). **Draft Standard for Learning Object Metadata**. IEEE-SA Standard 1484.12.1. New York: Institute of Electrical and Electronics Engineers, Inc.

IEEE. (2002). **Draft standard for learning Object Metadata**. [On-line]. Available: http://ltsc.ieee.org/ wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

IMS Global Learning Consortium. (2004). **IMS Metadata Best Practice Guide for IEEE 1484.12.1-2002 Standard for Learning Object Metadata**. [On-line]. Available: http://www.imsglobal.org /metadata/mdv1p3pd/imsmd_bestv1p3pd.html

Jansen, B. J. and Spink, A. (2003). An Analysis of Web Information Seeking and Use: Documents Retrieved Versus Documents Viewed. **Proceedings of the 4th International Conference on Internet Computing** (pp 339-355). New York: Association for Computing Machinery Press.

Jiang, J. J. and Conrath, D. W. (1997). Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. In **Proceeding of International Conference Research on Computational Linguistics** (pp 296-304). San Francisco: Kaufmann Publishers.

John F. and Paul S. (2003). **International encyclopedia of information and library science**. New York: Taylor & Francis Group.

Kalfoglu Y. and Schorlemmer M. (2003). Ontology mapping: the state of the art. In **The knowledge Engineering Review.** 18 (1): 1-31.

Kaufmann, E. and Bernstein, A. (2007). How useful are natural language interfaces to the semantic web for casual end-users? **Proceedings of the 6th international semantic web conference and 2nd Asian conference on Asian semantic web** (pp 281-294). New York: Springer Berlin Heidelberg.

Kayed, A. and Colomb, R. M. (2005). Using BWW model to evaluate building ontologies in CGs formalism. **Information Systems.** 30 (5): 379-398.

Keen, E.M. (1971). Evaluation parameters. In Gerard Salton (ed.). **The SMART retrieval system: experiments in automatic document processing** (pp 74-111). Englewood Cliffs, NJ: Prentice Hall.

Kent, A. (1955). Machine literature searching. VIII. Operational criteria for designing information retrieval systems. **American documentation**. 6 (2): 93-101.

Khoo, C. G., and Wan, K. -W. (2004). A simple relevancy-ranking strategy for an interface to Boolean OPACs. **The Electronic Library**. 22 (2): 112–120.

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. **Journal of the ACM.** 46 (5): 604-632.

Klyne, G. and Caroll, J. J. (2008). **Resource Description Framework (RDF): Concepts and Abstract Syntax**. [On-line]. Available: http://www.w3.org/TR/rdf-concepts/

Koll, M. (2000). Track 3: Information Retrieval. **Bulletin of the American society for information science**. 26 (2): 16-18.

Korfhage, R R. (1997). **Information storage and retrieval**. New York: John Wiley & Sons.

Kowalski, G. (1997). **Information retrieval system: Theory and implementation**. Boston: Kluwer Academic Publishers.

Krejcie, R. V. and Morgan, D. W. (1970). Determining sample size for research activities. **Educational and Psychological measurement**. 30 (3): 607-610.

Lancaster, F. W. (2003). **Indexing and abstracting in theory and practice**. 3nd edition. Champaign-Urbana, IL: University of Illinois, Graduate School of Library and Information Science.

Lange, D. B. (1998). **Programming and deploying java mobile agents with aglets**. Boston: Addison-Wesley Longman Publishing.

Large, A., Tedd, L. A. and Hartley, R. J. (1999). **Information seeking in the online age: Principles and practice**. London: Bowker-Saur.

Leckie, J., Pettigrew, E. & Sylvain, C. (1996). Modeling the information seeking of professionals: a general model derived from research on engineers, health care professionals and lawyers. **Library Quarterly**. 66 (2): 161-193.

Lesk, M. (1997). **Practical digital libraries: Books, bytes and bucks**. Sanfrancisco: Morgan Kaufmann.

Lewandowski, D. (2009). Ranking library materials. **Library Hi Tech**. 27 (4): 584–593.

Library of Congress. (1987). **Subject Headings**. 10th Ed. Washington: Library of Congress.

Lin, D. (1998). An information-theoretic definition of similarity. In **Proceeding of The 15th International Conference on Machine Learning** (pp 296-304). San Francisco: Morgan Kaufmann Publishers

Liu, Y. (2009). Learning to rank for information retrieval. **Foundations and Trends in Information Retrieval**. 3 (3): 225-331.

Liu, Y., Gao, B., Liu, T. Y., Zhang, Ma, Y., Z., He, S. and Li, H. (2008). BrowseRank: Letting web users vote for page importance. In **Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval** (pp 451-458). New York: Association for Computing Machinery Press.

Madhu, G., Govardhan, A. and Rajinikanth, T. (2011). Intelligent semantic web search engines: a brief survey. **International Journal of Web & Semantic Technology**.2 (1): 34–42

Manuel, L., Juan C., Vidal, E. O., Alberto, B. and Senén, B. (2012). Semantic Linking of Learning Object Repositories to DBpedia. **Educational Technology and Society**. 15 (4): 47-61.

Marchionini, G. (1995). **Information seeking in electronic environments**. New York: Cambridge University Press.

Maron, M. and Kuhns, J. (1960). On relevance, probabilistic indexing and information retrieval. **Journal of the Association for Computing Machinery Press**. 7 (3): 216-244.

McGuinness, D. and Harmelen, F. (2004). **OWL Web Ontology Language Overview**. [On-line]. Available: http://www.w3.org/ TR/owl-features/

McGuinness, D., Fikes, R., Rice, J., and Wilder, S. (2000). The Chimaera Ontology Environment. In **Proceedings of the 17th National Conference on Artificial Intelligence (AAAI)** (pp 1123-1124). Massachusetts: Association for the Advancement of Artificial Intelligence Press.

McKiernan, G. (2001). **Beyond bookmarks: Schemes for Organizing the Web**. [On-line]. Available: www.public.iastate.edu/CYBERSTACKS/CTW.htm

Meiss, M. R., Menczer, F., Fortunato, S., Flammini, A. and Vespignani, A. (2008). Ranking web sites with real user traffic. In **Proceedings of the 1st ACM International Conference on Web Search and Data Mining** (pp 65-76). New York: Association for Computing Machinery Press.

Miao, G., Tatemura, J., Hsiung, W.P., Sawires, A. and Moser, L. (2009). Extracting Data Records from the Web Using Tag Path Clustering. In **Proceedings of the 18th International Conference on World Wide Web (WWW'09)** (pp 981-990). New York: Association for Computing Machinery Press.

Mile, S. (2010). Mapping attribution metadata to the open provenance model. **Future Generation Computer System**. 27 (6): 806-811.

Navarro, G. and Baeza-Yates, R. (1997). Proximal nodes: A model to query document databases by content and structure. **ACM Transactions on information systems**. 15 (4): 400-435.

Neilsen, J. 2012. **Introduction to Usability**. [On-line]. Available: http://www.useit.com/alertbox /20030825.html

Nikolov, A., Uren, V. and Motta, E. (2008). De Roeck A. Integration of semantically annotated data by the KnoFuss architecture. **Proceedings of the 16th International Conference on**

Knowledge Engineering and Knowledge Management (pp 265-274). New York: Springer Berlin Heidelberg

Niles, I. and Pease, A. (2001). **Towards a standard upper ontology. Formal Ontology in Information Systems**. New York: Association for Computing Machinery Press.

Noia, T. D., Mirizzi, R., Claudio, O. V., Davide, R. and Zanker, M. (2012). Linked Open Data to support content-based recommender systems. In **Proceeding of The 8th International Conference on Semantic Systems** (pp 1-8). New York: Association for Computing Machinery Press.

Nonaka, I. and Takeuchi, H. (1995). **The knowledge-creating company: how Japanese companies create the dynamics of innovation**. New York: Oxford university press.

Noy N. and Musen, A. M. (1999). Smart: Automated Support for Ontology Merging and Alignment. In **Proceeding of the 12th Banff Workshop on Knowledge Acquisition, Modeling, and Management** (pp 1-24). New York: Association for Computing Machinery Press.

Noy, N. F. (2004). Semantic integration: A survey of ontology-based approaches. **ACM SIGMOD.** 33 (4): 65-70.

Noy, N. F. and Musen, M. A. (2004). Ontology versioning in an ontology management framework. **IEEE Intelligent System**. 19 (4): 6-13.

Nuno S. and Joao R. (2003). MAFRA – An Ontology Mapping FRAmework for the Semantic Web. In **Proceedings of the 6th International Conference on Business Information Systems; UCCS** (pp 1-9). Colorado Springs: UCCS Press.

O'Reilly, T. (2005). **What is web 2.0**. [On-line]. Available: oreilly.com/web2/archive/what-is/web-20.html

OECD, PISA (2003) **Learning for Tomorrow's World First Results from PISA 20.** [On-line]. Available: http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1_1, 0 0.html.

Oh, S. and Yeom, H. Y. (In Press). A comprehensive framework for the evaluation of ontology modularization. **Expert Systems with Applications**. [On-line]. Available: http://www.sciencedirect.com/science/article/pii/S0957417412001479

Orbst, L., Hughes, T. and Ray, S. (2006). Prospects and Possibilities for Ontology Evaluation: The View from NCOR. In **Proceedings of The 4th International Evaluation of Ontologies for the Web** (pp). New York: Springer Berlin Heidelberg.

Page, L., Brin, S., Motwani, R. and Winograd, T. (1999). **The PageRank citation ranking: Bringing order to the web**. Redwood City: Stanford University Press.

Park, J., Cho, W. and Rho, S. (2010). Evaluating ontology extraction tools using a comprehensive evaluation framework. **Data and Knowledge Engineering,** 69 (10): 1043-1061.

Pinto, H. S. and Joao P. M. (2001). A Methodology for Ontology Integration. in **Proceeding of the International Conference on Knowledge Capture, Technical papers** (pp 131-138). New York: Association for Computing Machinery Press.

Rahm, E. and Bernstein, P. A. (2001). A survey of approaches to automatic schema matching. **Very Large Database Journal.** 10 (4): 334-350.

Rao, A.S. and Georgeff, M. (1995). BDI Agents: from theory to practice. In **Proceeding of the 1st international conference on multi-agent systems** (pp 312-319). California: AAAI Press.

Raphael, V., Daniel, O., Steffen, S. and Boris, M. (2003). KAON SERVER A Semantic Web Management System. In **The 12th International World Wide Web Conference** (pp). New York: Association for Computing Machinery Press.

Resnik, P. (1995). Using Information Content to Evaluate Semantic Similarity in a Taxonomy. In **Proceedings of the 14th International Conference on Machine Learning** (pp 448-453). San Francisco: Morgan Kaufmann Publishers.

Roberson, S.E. (1977). The probability ranking principle in IR. **Journal of Documentation**. 33: 294-304.

Russell, S.J. and Norvig, P. (2003). **Artificial intelligence: a Modern Approach**. 2nd ed. New Jersey: Prentice-Hall.

Sahami, M., Yusufali, S. and Baldonado, W. (1998). SONIA: A Service for Organizing Networked Information Autonomously. In **Proceedings of The Digital Libraries** (pp 200-209). New York: Association for Computing Machinery Press.

Salton, G. (1968). **Automatic information organization and retrieval**. New York: McGraw-Hill.

Salton, G. (1971). **The SMART retrieval system: experiments in automatic document processing**. New Jersey: Prentice-hall.

Salton, G. (1992). The state of retrieval system evaluation. **Information processing & management**. 28 (4): 441-449.

Salton, G., McGill, M. J. (1983). **Introduction to Modern Retrieval**. New York: McGraw-Hill.

Samir, A., Ioan, M. B., Chabane, D. (2011). MuMIe: Multi-level Metadata Mapping System. **Journal of Multimedia**. 6 (3): 225-234.

Sbodio, M.L., Martin, D. and Moulin, C. (2010). Discovering Semantic Web services using SPARQL and intelligent agents, **Web Semantics: Science, Services and Agents on the World Wide Web**. 8 (4): 310-328,

Scott, M. L. (2005). **Dewey decimals classification, 22nd edition: A study manual and number building guide**. Westport: Libraries Unlimited.

Seaborne, Andy. (2004). **RDQL - A Query Language for RDF**. [On-line]. Available: http:// www.w3.org/Submission/2004/SUBM-RDQL-20040109/

Sears, M. E. (1977). **Sears List of Subject Headings**. 11th Ed. New York: H.W. Wilson.

Seksun Suwanmanee, Djamal Benslimane and Philippe Thiran. (2005). OWL-Based approach for semantic interoperability. In **Proceeding of the 19th International Conference on Advanced information networking and Applications** (pp 145-150). New Jersey: Institute of Electrical and Electronics Engineers.

Senge, P.M. (1990). **The fifth discipline: the art and practice of learning organization**. London: Century press.

Shatkay, H. and Wibur, W. J. (2000). Finding Themes in Medline Documents Probabilistic Similarity Search. In **Proceeding of Advances in Digital Libraries** (pp 183-192). New Jersey: Institute of Electrical and Electronics Engineers.

Shishechi S. S., Banihashem, S. Y., Zin, N. M. and Mohd, S. A. (2012). Ontological approach in knowledge based recommender system to develop the quality of e-learning system. **Australian journal of basic and applied sciences**. 6 (2): 115-123.

Shneiderman, B., Plaisant, C., Cohen, M. and Jacobs, S. (2009). **Designing the user interface: strategies for effective human- computer interaction**. 5th ed. Boston: Addison Wesley.

Skyrme, D. J. (2002). **The 3 Cs of knowledge sharing: Culture, co-operation and commitment**. [On-line]. Available: http://www.skyrme.com/updates/u64_f1.htm.

Smith, G. (2008). Tagging: Emerging trends. **Bulletin of the American Society for Information Science and Technology**. 43 (6): 1449-1481

Sofia, P. H., Gomez-Perez A. and Martins P. J. (1999). Some Issues on Ontology Integration. In **Proceeding of 16th International Joint Conference on Artificial Intelligence Workshop on Ontologies and Problem Solving Methods: Lessons Learned and Future Trends** (pp 7-1-7-12). California: AAAI Press.

Song, R., Liu, H., Wen, R. and Ma, Y. (2004). Learning important models for web page blocks based on layout and content analysis. **ACM SIGKDD Explorations Newsletter.** 6 (2): 14-23.

Sowa, J. F. (2000). **Knowledge Representation: Logical, Philosophical, and Computational Foundations.** Pacific Grove, CA: Brooks Cole.

Soysal, E., Cicekli, I. and Baykal, N. (2010). Design and evaluation of an ontology based information extraction system for radiological reports. **Computers in Biology and Medicine,** 40 (11–12): 900-911.

Sparck, J. K. (2000). Further reflections on TREC. **Information processing & management**. 36 (1): 37-85.

Sparck, J. K. and Willett, P. (1997). **Readings in information retrieval**. San Francisco: Morgan Kaufmann.

Stielow, F. nd Tibbo, H. (1988). The negative search, on-line reference, and the humanities: A critical essay in library literature. **RQ**. 27 (3): 358-365.

Strasberg, H. R., Manning, C. D., Rindfleisch, T. C. and Melmon, K. L. (2000). What's Related? Generalizing Approaches to Related Articles in Medicine. In **Proceeding of AMIA Symposium** (pp 838-842). Philadelphia: Hanley & Belfus.

Sullivan, D. (2015). **The periodic table of SEO success factors**. [On-line]. Available: http://search engineland.com/seotable

Sure, Y., Gomez-Perez, G.-P., Daelemans, W., Reinberger, M.-L., Guarino, N. and Noy, N. F. (2004). Why evaluate ontology technologies?: Because it works. **IEEE Intelligent Systems,** 19 (4): 74-81.

Surowiecki, J. (2005). **The wisdom of crowds. New York**. USA: Anchor Books.

Swartout, W. R. and Tate, A. (1999). Guest editors' introduction: Ontologies. **IEEE Intelligent Systems,** 14 (1): 18–19.

Termehchy, A. and Winslett, M. (2010). Keyword search over key-value stores. In **Proceedings of the 19th International Conference on World Wide Web** (pp 1193-1194). New York: Association for Computing Machinery Press.

Thornely, J. (1998). The Road To Meta: The Implementation of Dublin Core Metadata in The State Library of Queensland Website. **The Australian Library journal**. 47 (1): 74-82.

TIMSS International Study Center Boston College Chestnut Hill (2011) **Highlights from TIMSS 2011: Mathematics and Science Achievement of U.S. Fourth and Eighth-Grade Students in an International Context**. [On-Line]. Available: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2013009rev

Toupikov, N., Umbrich, J., Delbru, R., Hausenblas, M. and Tummerello, G. (2009). Ding! Dataset Ranking using formal descriptions. In **Proceeding of Linked Data On the Web** (pp 201-205). New Haven: CEUR-Workshop Proceeding Group.

Tsoi, L. C., Patel, R., Zhao, W., and Zheng, W. J. (2009). Text-mining approach to evaluate terms for ontology development. **Journal of Biomedical Informatics,** 42 (5): 824-830.

Uschold, M. and Gruninger M. (1996). Ontologies: Principles, Methods and Application. **Knowledge Engineering Review**. 11(2): 93-136.

Vander, W. T. (2007). **Folksonomy**. [On-line]. Available: vanderwal.net/folksonomy.html

Wang, Z., Juanzi, L., Zhao, Y., Setchi, R. and Tang, J. (2013). A Unified approach to matching semantic data on the web. **Knowledge-based systems**. 39(3): 173-184.

Weibel, S. (1997). The Dublin Core: A Simple Content Description Model for Electronic Resources. **Bulletin of The American Society for Information Science**. 24 (1): 9-11.

Weideman, M. (2009). **Website visibility: The theory and practice of improving rankings**. Oxford; Cambridge. New Delhi: Chandos Publishing.

Wellisch, H. H. (1995). **Indexing from A to Z**. 2nd edition. New York: H.W. Willson.

Wiley, D. A. (2000). **Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy**. Utah: Utah State University.

Wiley, D. A. and Nelson, L. M. (1998). **The fundamental object**. [On-line]. Available: http://wiley.ed.usu.edu/docs/fundamental.html

Wilson, T. D. (1999). Models in information behaviour research. **Journal of Documentation**, 55 (3): 249-270.

Wilson, T. D. (2000). **Human information behavior** (Tech. Rep. No.2). Sheffield: University of Sheffield, Information Science Research.

Wirth, D. and Brooks J. (2006). **An Introduction to the Advanced Distributed Learning Initiative**. [On-line]. Available: http://www.adlnet.gov.

Wooldridge, M. J. and Jenings, N.R. (1995). Intelligent agents: theory and practice. **Knowledge engineering review**. 10 (2): 115-152.

Wu, Z. and Palmer, M. (1994). Verb semantics and lexical selection. In **Proceedings of the 32nd Annual meeting of the Association for Computational Linguistics, Las Cruces** (pp 133-138). Stroudsburg: Association for Computational Linguistics.

Wynar, B. S. (1980). **Introduction to cataloguing and classification 6th ed**. CO: Libraries Limited.

Xu, C. and Chu, H. (2008). Social tagging in China and the USA: a comparative study. In **Proceeding of the annual meeting of the American society for information science and technology** (pp 1-9). Mary Land: ASIS&T.

Yoosooka, B. and Wuwongse, V. (2011). Linked Open Data for learning object discovery. In **Proceeding of the International Conference on Intelligent Networking and Collaborative System** (pp 60-67). New York: Institute of Electrical and Electronics Engineers.

Zhao, L. and Ichise, R. (2012). Graph-based ontology analysis in the linked open data. In **Proceeding of The 8th conference on semantic systems** (pp 56-63). New York: Association for Computing Machinery Press.

Zhou, D., Lawless, S., Min, J. and Wade, V. (2010). A late fusion approach to cross-lingual document re-ranking. **In Proceedings of the 19th ACM international conference on Information and knowledge management** (pp 1433-1436). New York: Association for Computing Machinery Press.

ภาคผนวก ก.

แบบสำรวจเพื่อการวิจัย

แบบสอบถามการค้นคืนและเลือกใช้ข้อมูลจากแหล่งสารสนเทศบนอินเทอร์เน็ต

มหาวิทยาลัยเทคโนโลยีสุรนารี

แบบสอบถามการค้นคืนและเลือกใช้ข้อมูลจากแหล่งสารสนเทศบนอินเทอร์เน็ต เพื่อใช้สร้างสื่อการเรียนรู้ประกอบการเรียนการสอนในรายวิชาคณิตศาสตร์ในหลักสูตรสองภาษา ระดับมัธยมศึกษาชั้นปีที่ 1-6 ตามหลักสูตรแกนกลางขั้นพื้นฐาน พุทธศักราช 2551

คำชี้แจง แบบสอบถามครูผู้สอนรายวิชาคณิตศาสตร์ในหลักสูตรสองภาษา ระดับมัธยมศึกษาชั้นปีที่ 1 - 6

เรื่อง การค้นคืนและเลือกใช้ข้อมูลเนื้อหารายวิชาคณิตศาสตร์จากแหล่งสารสนเทศบนอินเทอร์เน็ต เพื่อใช้ สร้างสื่อการเรียนรู้ประกอบการเรียนการสอน ตามหลักสูตรแกนกลางขั้นพื้นฐาน พุทธศักราช 2551

แบบสอบถามฉบับนี้เป็นส่วนหนึ่งของการทำวิทยานิพนธ์ในระดับปริญญาเอก สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี โดยมีวัตถุประสงค์เพื่อศึกษาถึงปัจจัยที่ครูผู้สอนใช้ในการพิจารณาคัดเลือกเนื้อหาที่เกี่ยวกับรายวิชาคณิตศาสตร์ที่ค้นคืนได้จากแหล่งสารสนเทศบนอินเทอร์เน็ต เพื่อนำมาใช้สร้างสื่อการเรียนรู้แบบฝึกหัด งานมอบหมาย ตลอดจนใช้เป็นเนื้อหาประกอบในการเรียนการสอนของครูผู้สอนรายวิชาคณิตศาสตร์ ในหลักสูตรสองภาษา ระดับมัธยมศึกษาชั้นปีที่ 1-6 ผู้วิจัยคาดหวังอย่างยิ่งว่าผลการวิเคราะห์แบบสอบถามครั้งนี้จะเป็นประโยชน์ในการนำไปประกอบกระบวนการพัฒนาเครื่องมือในการค้นคืนข้อมูลเพื่อจัดทำสื่อการเรียนรู้ใช้ในการเรียนการสอนที่มีประสิทธิภาพมากขึ้น โดยศึกษาอ้างอิงวิธีการค้นคืนและปัจจัยการพิจารณาคัดเลือกข้อมูลจากผู้ใช้เป็นสำคัญ ทั้งนี้ผู้วิจัยจะเก็บข้อมูลของท่านไว้เป็นความลับ โดยจะนำไปใช้เพื่อสรุปผลการวิจัยในภาพรวมเกี่ยวกับวิธีการค้นคืนและการคัดเลือกข้อมูลเท่านั้น

ดังนั้นการที่ท่านให้ข้อมูลที่ตรงกับความเป็นจริงและสมบูรณ์จะช่วยให้การวิจัยดำเนินไปด้วยความถูกต้อง ผู้วิจัยจึงใคร่ขอความอนุเคราะห์จากท่าน โปรดตอบแบบสอบถามตามความคิดเห็นและตามการใช้งานจริงของท่านให้ครบทุกข้อ และขอความกรุณาจัดส่งเอกสารกลับมาทางซองที่จัดเตรียมไว้ให้กับแบบสอบถามนี้ ถ้าท่านมีข้อสงสัยประการใด สามารถติดต่อกลับทางผู้วิจัยทางหมายเลขโทรศัพท์ตอนท้ายของแบบสอบถามได้

แบบสอบถามชุดนี้แบ่งออกเป็น 3 ตอน จำนวน 6 หน้า

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถามเพื่อใช้ในการอ้างอิงแหล่งข้อมูลในการวิจัย

ตอนที่ 2 แบบสอบถามการเลือกใช้เครื่องมือและแหล่งค้นคืนสารสนเทศบนอินเทอร์เน็ต เพื่อค้นหาข้อมูลเกี่ยวกับรายวิชาคณิตศาสตร์

ตอนที่ 3 แบบสอบถามการเลือกใช้ข้อมูลจากแหล่งสารสนเทศ เพื่อนำมาใช้ประกอบการเรียนการสอนรายวิชาคณิตศาสตร์

ตอนที่ 4 แบบสอบถามสภาพปัจจุบันในการจัดเก็บเลิร์นนิงอ็อบเจกต์ และลักษณะความต้องการระบบการจัดเก็บเลิร์นนิงอ็อบเจกต์

ขอขอบพระคุณในความกรุณาของท่าน

นายนพพล ตั้งสุภาชัย

สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม

มหาวิทยาลัยเทคโนโลยีสุรนารี

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตอนที่ 1 ข้อมูลสภาพทั่วไปของผู้ตอบแบบสอบถาม

คำชี้แจง: กรุณาใส่เครื่องหมาย ✓ ลงใน ❑ หน้าคำตอบที่ตรงกับสภาพความเป็นจริงและกรุณากรอกรายละเอียดลงในช่องว่างที่กำหนดถ้าเลือกตัวเลือกในข้อนั้นๆ

1. สถานภาพของท่าน ❑ อาจารย์ผู้สอน ❑ อาจารย์ผู้ฝึกสอน ❑ ผู้ช่วยสอน
2. ตำแหน่งทางวิชาการ/ระดับอาจารย์_____
3. ประเภทของโรงเรียนที่สังกัด ❑ โรงเรียนรัฐบาล ❑ โรงเรียนเอกชน
4. ประสบการณ์ในการสอน ❑ 1-3 ปี ❑ 4-6 ปี ❑ 7-9 ปี ❑ 10 ปีขึ้นไป
5. ชั้นปีการศึกษาที่สอนคณิตศาสตร์ (ตอบได้มากกว่า 1 ข้อ)
   ❑ ชั้น ม.1 ❑ ชั้น ม.2 ❑ ชั้น ม.3 ❑ ชั้น ม.4 ❑ ชั้น ม.5 ❑ ชั้น ม.6
6. เนื้อหาคณิตศาสตร์ที่สอนเป็น (ตอบได้มากกว่า 1 ข้อ)
   ❑ ภาคภาษาไทยเท่านั้น ❑ ภาคภาษาอังกฤษเท่านั้น ❑ ทั้งสองภาษา

ตอนที่ 2 แบบสอบถามการเลือกใช้เครื่องมือและแหล่งค้นคืนสารสนเทศบนอินเทอร์เน็ต เพื่อค้นหาข้อมูลเกี่ยวกับรายวิชาคณิตศาสตร์

คำชี้แจง: กรุณาใส่เครื่องหมาย ✓ ลงใน ❑ หน้าคำตอบที่ตรงกับสภาพความเป็นจริง และกรุณากรอกลำดับตัวเลขเพื่อลงรายละเอียดในช่องว่างที่กำหนดถ้าเลือกตัวเลือกในข้อนั้นๆ

1. ท่านรู้จักและเคยใช้งานอินเทอร์เน็ตมานานเพียงใด

❑ น้อยกว่า 6 เดือน ❑ 6 เดือน ถึง 1 ปี
❑ 1ปี ถึง 2 ปี ❑ 2 ถึง 3 ปี
❑ 3 ถึง 4 ปี ❑ มากกว่า 4 ปี

2. สถานที่ที่ท่านใช้อินเทอร์เน็ตบ่อยที่สุดคือสถานที่ใด

❑ ที่พักส่วนตัว, ที่บ้าน ❑ คอมพิวเตอร์ที่ห้องทำงาน, ห้องพักครู
❑ ที่ห้องสมุดของโรงเรียน ❑ ร้านอินเทอร์เน็ตภายนอก
❑ อื่นๆ โปรดระบุ_____

3. ความถี่ของการใช้งานอินเทอร์เน็ตของท่านบ่อยเพียงใด

❑ 1-2 ครั้ง/สัปดาห์ ❑ 3-4 ครั้ง/สัปดาห์
❑ 5-6 ครั้ง/สัปดาห์ ❑ มากกว่า 6 ครั้งต่อสัปดาห์

มหาวิทยาลัยเทคโนโลยีสุรนารี

4. ระยะเวลาโดยเฉลี่ยที่ใช้งานอินเทอร์เน็ตต่อครั้งนานเท่าใด

❑ น้อยกว่า 1 ชั่วโมง ❑ 1-2 ชั่วโมง

❑ 2-4 ชั่วโมง ❑ มากกว่า 4 ชั่วโมง

5. แหล่งข้อมูลที่ท่านใช้เพื่อค้นหาข้อมูลและใช้ประกอบสื่อการเรียนการสอนในวิชาคณิตศาสตร์คือที่ใด (ตอบได้มากกว่า 1 ข้อ)

❑ หนังสือเรียนหลักตามกำหนดของหลักสูตรแกนกลางขั้นพื้นฐานของ สพฐ.

❑ หนังสือเรียนเสริมจากที่โรงเรียนจัดหาเพิ่ม

❑ หนังสือเรียนเสริมจากสำนักพิมพ์เอกชนทั่วไป

❑ สื่อการเรียนการสอนเสริมอื่น ๆ เช่น CD ช่วยสอน

❑ สื่อการเรียนรู้หรือเนื้อหาจากอินเทอร์เน็ต

❑ อื่นๆ โปรดระบุ_____

6. เครื่องมือหรือวิธีที่ท่านนิยมใช้เพื่อเข้าถึงข้อมูลบนอินเทอร์เน็ตได้มาจากที่ใด (ตอบเรียงลำดับจาก มาก ไป น้อย โดยใช้ตัวเลข 1 นิยมใช้มากที่สุด ถึง 5 นิยมใช้น้อยที่สุด)

_____ค้นหาจากเว็บค้นคืนข้อมูล (Search engine) เช่น Google, Yahoo, Lycos เป็นต้น

_____ค้นหาจากเว็บจัดกลุ่มข้อมูล (Web directory) เช่น Yahoo, sanook, kapook เป็นต้น

_____เข้าเว็บของหน่วยงานราชการที่เกี่ยวข้อง เช่น เว็บ สวทช, สสวท หรือเว็บของกระทรวงต่างๆ

_____เข้าเว็บคลังข้อมูลเกี่ยวกับการเรียนการสอน เช่น Schoolnet, Thaigoodview เป็นต้น

_____อื่นๆ โปรดระบุ_____

7. วิธีการใดที่ท่านนิยมใช้ในการค้นคืนข้อมูลบนอินเทอร์เน็ต (ตอบเรียงลำดับจาก มาก ไป น้อย โดยใช้ตัวเลข 1 นิยมใช้มากที่สุด ถึง 5 นิยมใช้น้อยที่สุด)

_____ค้นจากคำสำคัญ (Keyword)

_____ค้นจากชื่อบุคคลที่เกี่ยวข้องกับเนื้อหา

_____ค้นจากชื่อเรื่อง หรือชื่อบทเรียน

_____ค้นจากชื่อหนังสือ หรือเอกสารการสอน

_____ค้นจากหลายเขตข้อมูลรวมกัน เช่น ค้นหาจากชื่อบุคคล, ชื่อเรื่อง และคำสำคัญ พร้อมๆ กัน

_____อื่นๆ โปรดระบุ_____

8. ลักษณะของข้อมูลเกี่ยวกับวิชาคณิตศาสตร์ที่ท่านมักจะค้นหาเป็นข้อมูลลักษณะใด (ตอบเรียงลำดับจาก มาก ไป น้อย โดยใช้ตัวเลข 1 ใช้มากที่สุด ถึง 5 ใช้น้อยที่สุด)

_____เนื้อหาบรรยายเกี่ยวกับเรื่องที่สอน

_____โจทย์แบบฝึกหัดก่อนเรียน และหลังเรียน

_____ข้อสอบหรือแบบทดสอบวัดผลการเรียน

_____รูปภาพประกอบเนื้อหารายวิชา

_____กิจกรรมหรือเกมที่เกี่ยวกับเนื้อหารายวิชา

_____อื่นๆ โปรดระบุ_____

9. ข้อมูลที่ค้นคืนได้จากอินเทอร์เน็ตท่านนำไปใช้เกี่ยวกับการเรียนการสอนวิชาคณิตศาสตร์อย่างไร (ตอบเรียงลำดับจาก มาก ไป น้อย โดยใช้ตัวเลข 1 ใช้มากที่สุด ถึง 5 ใช้น้อยที่สุด)

_____ใช้ประกอบสื่อการเรียนการสอนหลัก

_____ใช้เป็นส่วนเสริมในเนื้อหาการเรียนการสอนเพื่อให้มีความหลากหลาย

_____ใช้ศึกษาเพิ่มเติมความรู้ในเรื่องใหม่ๆ ของท่านเองเพื่อเทคนิคการสอนที่ดีขึ้น

_____ใช้เพื่อเขียนตำราหรือหนังสืออ้างอิง

_____ใช้เพื่อเป็นแบบฝึกหัดเสริม หรือการเรียนเสริมนอกเวลาเรียนของนักเรียน

_____อื่นๆ โปรดระบุ_____

10.กรุณาใส่เครื่องหมาย ✓ ลงในตารางแสดงระดับการพิจารณาความสำคัญของปัญหาที่ท่านพบจากการค้นคืนข้อมูลบนอินเทอร์เน็ต

ปัญหาที่เกิดจากการค้นคืนข้อมูล	ระดับความสำคัญของปัญหา				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
1. ด้านทักษะในการค้นคืน					
1.1 ขาดความชำนาญในการสืบค้น					
1.2 ขาดทักษะภาษาอังกฤษเพื่อสืบค้น					
1.3 ขาดความรู้การเลือกใช้คำค้นที่เหมาะสม					
1.4 ขาดความรู้เกี่ยวกับแหล่งที่ใช้สืบค้น					

2. ด้านรายละเอียดของผลการค้นคืน					
2.1 ผลลัพธ์มีความซ้ำซ้อนหรือเหมือนกันในหลายรายการของผลการสืบค้น					
2.2 ชื่อเรื่องไม่สอดคล้องกับเนื้อหาหรือเป็นการหลอกล่อเพื่อเข้าไปยังเว็บที่เป็นเนื้อหาอื่นๆ					
2.3 ข้อมูลมีเนื้อหาไม่ถูกต้อง					
2.4 รายชื่อสื่อ (ชื่อเรื่อง) ที่เป็นผลลัพธ์การค้นคืนไม่ปรากฏคำค้นที่ใช้ ทำให้มีผลกับการตัดสินใจเลือกใช้ผลการค้นคืนแต่ละรายการ					
2.5 คำอธิบายของผลลัพธ์ (บทคัดย่อ) แต่ละรายการไม่ปรากฏคำค้นที่ใช้ ทำให้มีผลกับการตัดสินใจเลือกใช้ผลการค้นคืนแต่ละรายการ					
2.6 รายการผลลัพธ์มีจำนวนที่มากเกินไป					
2.7 ไม่มีการแนะนำคำค้นที่เกี่ยวข้อง					
2.8 ไม่มีการเชื่อมโยงไปยังเว็บไซต์หรือแหล่งข้อมูลอื่นที่เกี่ยวข้องเพิ่มเติม					
2.9 ไม่มีคำวิจารณ์หรือการให้คะแนนรายการผลการค้นคืนแต่ละรายการ					
2.10 ไม่ระบุแหล่งที่มาหรือดูไม่น่าเชื่อถือ					
2.11 ข้อมูลมีลิขสิทธิ์หรือมีค่าใช้จ่าย					

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตอนที่ 3 แบบสอบถามการเลือกใช้ข้อมูลจากแหล่งสารสนเทศ เพื่อนำมาใช้ในการสร้างสื่อการเรียนรู้ และนำไปประกอบการเรียนการสอนรายวิชาคณิตศาสตร์ หลักสูตรสองภาษา

คำชี้แจง: กรุณาใส่เครื่องหมาย ✓ ลงในตารางแสดงระดับการพิจารณาความสำคัญของข้อมูลต่างๆ ที่ใช้ในการเลือกใช้สารสนเทศบนอินเทอร์เน็ต

ข้อมูลที่ใช้ในการพิจารณาคัดเลือก "รายการผลการค้นคืนบนอินเทอร์เน็ต"

ปัจจัยการพิจารณา	ระดับการให้ความสำคัญในการพิจารณา				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
1. ชื่อเรื่องเต็ม (Title)					
2. คำสำคัญที่ปรากฏบนชื่อเรื่อง (Keyword)					
3. รายละเอียดเนื้อหา ส่วนประกอบภายในชิ้นงาน (Abstract/ Description)					
4. ชื่อผู้แต่งหรือผู้สร้างผลงาน (Author)					
5. หัวเรื่อง/ประเด็นหลักของเนื้อหา (Subject Heading)					
6. รหัสระบุตามมาตรฐานสากล เช่น URI ISBN เป็นต้น					
7. ชื่อหน่วยงานหรือเว็บไซต์ที่เผยแพร่ข้อมูล (Publisher)					
8. ภาษาที่ใช้ในชิ้นงาน (Language)					
9. แหล่งที่มาของข้อมูลในเนื้อหา (Source)					
10. ข้อมูลแสดงสิทธิ์ในการเข้าใช้ หรือแสดงลิขสิทธิ์ (Copyright)					
11. วันที่/ปีที่ผลิตและเผยแพร่เอกสาร (Date)					
12. แหล่งข้อมูลเชื่อมโยงอื่นๆ (Related)					
13. ระดับความยากง่ายของเนื้อหา					

มหาวิทยาลัยเทคโนโลยีสุรนารี

14. ประเภทของไฟล์ (Word, PDF, Excel เป็นต้น)					
15. ลำดับการแสดงผลรายการก่อน-หลัง					
16. คำวิจารณ์หรือการให้คะแนน (Review)					

17. ท่านพิจารณาความ ยาก-ง่าย ของเนื้อหาว่าเหมาะสมกับระดับที่ท่านสอนอย่างไร

_____

_____

_____

18. วิธีการคัดเลือกรายการผลค้นคืนแบบอื่นๆ ที่ท่านเลือกใช้ โปรดระบุ

_____

_____

วิธีการคัดกรองหรือปรับแต่งการแสดงผลการสืบค้น

วิธีการคัดกรองข้อมูล	รูปแบบการแสดงผลที่ท่านต้องการ				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
19. แสดงเฉพาะรายการที่เป็นฉบับเต็ม (Full text)					
20. การจัดกลุ่มเนื้อหาตามระดับชั้นปีการศึกษา (Level)					
21. การจัดกลุ่มเนื้อหาตามประเด็นเรื่อง (Subject)					
22. แสดงเฉพาะประเภทของชิ้นงานที่ต้องการ เช่น เป็นแบบฝึกหัด เอกสารบรรยายเนื้อหา ข้อสอบ การบ้าน เป็นต้น					
23. การแสดงรูปภาพหรือสื่ออื่นๆ ที่อยู่ในเนื้อหาออกมาไว้ในรายการผลการค้นคืน					

มหาวิทยาลัยเทคโนโลยีสุรนารี

ตอนที่ 4 แบบสอบถามสภาพปัจจุบันในการจัดเก็บเลิร์นนิงอ็อบเจกต์ และลักษณะความต้องการระบบการจัดเก็บเลิร์นนิงอ็อบเจกต์

คำชี้แจง: กรุณาใส่ลำดับคะแนนหน้าข้อแสดงระดับการพิจารณาความสำคัญของข้อมูลต่างๆ ที่ใช้ในการจัดเก็บเลิร์นนิงอ็อบเจกต์

เมื่อท่านสร้างสื่อการเรียนรู้พร้อมใช้งานแล้ว ท่านจัดเก็บสื่อการเรียนรู้อย่างไร (ตอบเรียงลำดับจากมาก ไป น้อย โดยใช้ตัวเลข 1 ทำมากที่สุด ถึง 5 น้อยที่สุด)

_____จัดเก็บในคอมพิวเตอร์ส่วนตัวไม่มีการเผยแพร่

_____ใช้โปรแกรมที่สถานบันกำหนด CMS หรือคลังข้อมูลส่วนกลาง

_____ใช้โปรแกรมสำเร็จรูปในการจัดเก็บเป็นระบบส่วนตัว

_____ไม่จัดเก็บเป็นไฟล์ แต่พิมพ์เอกสารออกมาเพื่อใช้งาน

_____อื่นๆ โปรดระบุ_____

ข้อเสนอแนะอื่นๆ

_____

_____

_____

ขอขอบพระคุณที่กรุณาตอบแบบสอบถาม

ผู้วิจัยในหลักสูตรปริญญาเอกสาขาวิชาเทคโนโลยีสารสนเทศ

สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม

มหาวิทยาลัยเทคโนโลยีสุรนารี

หมายเลขสำหรับติดต่อ : นายนพพล ตั้งสุภาชัย โทร.086-0

มหาวิทยาลัยเทคโนโลยีสุรนารี

ภาคผนวก ข.

แบบประเมินประสิทธิภาพแบบจำลองการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์

แบบประเมินการทำงานของแบบจำลองการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์
การวิจัยเรื่อง การพัฒนาแบบจำลองการจัดเก็บและค้นคืนเลิร์นนิงอ็อบเจกต์
ในชุดข้อมูลจากลิงค์โอเพนดาทา

.....

คำชี้แจง

1. แบบสอบถามชุดนี้ แบ่งออกเป็น 2 ตอน ประกอบด้วย

ตอนที่ 1 ข้อมูลทั่วไปของผู้ประเมินระบบ จำนวน 2 ข้อ

ตอนที่ 2 ความคิดเห็นเกี่ยวกับการทำงานของแบบจำลองการจัดเก็บและค้นคืนเลิร์นนิง-อ็อบเจกต์ จำนวน 4 ด้าน รวม 20 ข้อ

2. การตอบแบบสอบถามนี้ คำตอบของท่านมีความสำคัญอย่างยิ่งต่อการวิเคราะห์ข้อมูลที่ใช้ในการวิจัย กรุณาตอบข้อมูลให้ครบถ้วนทุกข้อ ทุกตอน เพื่อให้การวิจัยในครั้งนี้มีความเที่ยงตรงและเกิดประโยชน์อย่างแท้จริง

3. ข้อมูลที่ท่านตอบแบบสอบถามนี้จะเก็บไว้เป็นความลับ การนำเสนอผลการวิจัยจะเสนอในภาพรวมเท่านั้น

มหาวิทยาลัยเทคโนโลยีสุรนารี

ขอขอบพระคุณที่ให้ข้อมูลไว้ ณ โอกาสนี้

นายนพพล ตั้งสุภาชัย
นักศึกษาระดับปริญญาเอก
สาขาวิชาเทคโนโลยีสารสนเทศ สานักวิชาเทคโนโลยีสังคม
มหาวิทยาลัยเทคโนโลยีสุรนารี

คำชี้แจง: กรุณาใส่เครื่องหมาย ✓ ลงใน ❑ หน้าคำตอบที่ตรงกับสภาพความเป็นจริงและกรุณากรอกรายละเอียดลงในช่องว่างที่กำหนดถ้าเลือกตัวเลือกในข้อนั้นๆ

ตอนที่ 1 ข้อมูลสภาพทั่วไปของผู้ตอบแบบสอบถาม

1.1 เพศ ❑ชาย ❑หญิง

1.2 ระดับการศึกษาสูงสุด

❑ ต่ำกว่าปริญญาตรี ❑ ปริญญาตรีหรือเทียบเท่า

❑ ปริญญาโท ❑ ปริญญาเอก

ตอนที่ 2 แบบสอบถามเกี่ยวกับประสิทธิภาพของระบบ

รายการคำถาม	ระดับความคิดเห็น				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
1.มอดูลการจัดเก็บข้อมูลเลิร์นนิงอ็อบเจกต์					
1.1 เมนูการเพิ่มข้อมูลสื่อความหมายและเข้าถึงได้ง่าย					
1.2 ฟอร์มกรอกข้อมูลมีเขตข้อมูลเหมาะสมชัดเจน					
1.3 ลำดับการเพิ่มข้อมูลมีขั้นตอนที่เข้าใจง่าย					
1.4 ข้อมูลที่เพิ่มเสร็จมีความครบถ้วนตามต้นฉบับ					
2 มอดูลการนำเข้าข้อมูลเลิร์นนิงอ็อบเจกต์					
2.1 ความสะดวกในการเรียกใช้เมนูคำสั่ง					
2.2 กระบวนการนำเข้ามีขั้นตอนที่ง่ายไม่ยุ่งยาก					
2.3 ระบบแจ้งขั้นตอนการดำเนินงานเข้าใจง่าย					

มหาวิทยาลัยเทคโนโลยีสุรนารี

2.4 ข้อมูลมีความสมบูรณ์ครบถ้วนตามต้นฉบับ					
3 มอดูลการค้นคืนเลิร์นนิงอ็อบเจต์					
3.1 ช่องกรอกคำค้นมีตัวเลือกการค้นคืนเหมาะสม					
3.2 การค้นคืนแบบไร่เรียงมีเครื่องมือค้นที่เหมาะสม					
3.3 ผลการค้นหาแบบแบ่ง 2 ระดับมีความเหมาะสม					
3.4 องค์ประกอบผลการค้นคืนเป็นประโยชน์ต่อการเลือกใช้					
3.5 การจัดกลุ่มเนื้อหาที่เกี่ยวข้องมีความเหมาะสม					
3.6 ระยะเวลาในการรอผลการค้นคืน					
3.7 ผลการค้นคืนตรงกับความต้องการของผู้ใช้					
4 ความง่ายในการใช้งาน					
4.1 ขนาด สี รูปแบบอักษร มีความเหมาะสม					
4.2 การออกแบบกราฟิก มีความชัดเจนเหมาะสม					
4.3 การจัดวางองค์ประกอบของระบบดูเหมาะสม					
4.4 แนวทางการเชื่อมโยงไปผลการค้นคืน และดาวน์โหลดเอกสาร มีความเหมาะสม					
4.5 รูปแบบของการระบุคำค้นสะดวกในการใช้งาน					
4.6 ความเหมาะสมในการปฏิสัมพันธ์กับผู้ใช้					

มหาวิทยาลัยเทคโนโลยีสุรนารี

ข้อเสนอแนะและความคิดเห็นอื่นๆ

……

……

……

(ลงชื่อ).....ผู้ประเมิน

(.....)

ความคิดเห็นของผู้ใช้ คนที่.....

มหาวิทยาลัยเทคโนโลยีสุรนารี

ประวัติผู้วิจัย

นายนพพล ตั้งสุภาชัย เกิดเมื่อวันเสาร์ที่ 16 มกราคม พ.ศ.2525 ภูมิเลานำเป็นคนจังหวัดนครราชสีมาโดยกำเนิด สำเร็จการศึกษาระดับปริญญาตรีในหลักสูตรวิทยาการสารสนเทศบัณฑิต ในกลุ่มวิชาสารสนเทศศึกษา สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี ในปี พ.ศ. 2546 และได้สำเร็จการศึกษาระดับปริญญาโท ในคณะวิทยาศาสตรมหาบัณฑิต คณะวิชาเทคโนโลยีสารสนเทศ จากสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ในปี พ.ศ.2549 จากนั้นได้เข้าปฏิบัติงานในตำแหน่งเจ้าหน้าที่วิเคราะห์และพัฒนาระบบ ที่บริษัท เซ็นทรัลพัฒนา จำกัด (CPN) กรุงเทพมหานคร จากนั้นได้ย้ายกลับภูมิลำเนา และได้เข้าทำงานที่ มหาวิทยาลัยวงษ์ชวลิตกุลนครราชสีมา ในตำแหน่งอาจารย์ประจำ สาขาวิชาคอมพิวเตอร์ธุรกิจ คณะวิชาบริหารธุรกิจ และในปี พ.ศ. 2553 ได้เข้าศึกษาต่อในระดับปริญญาเอก ในหลักสูตรวิทยาการสารสนเทศดุษฎีบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี ได้รับทุนสนับสนุนการศึกษาประเภท ทุนการศึกษาแก่นักศึกษาระดับบัณฑิตศึกษาที่คณาจารย์ได้รับทุนวิจัยจากแหล่งทุนภายนอก (OROG) โดยอาจารย์ ดร.นิศาชล จำนงศรี อาจารย์ประจำสาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี ทั้งนี้ ผู้วิจัยได้ศึกษาและมีความสนใจในด้านการจัดการความรู้ การออกแบบและสร้างออนโทโลยี และการวิจัยทางด้านระบบห้องสมุดอัตโนมัติ และห้องสมุดดิจิทัล การออกแบบระบบฐานข้อมูลและการผสานเมทาดาทา รวมถึงการประยุกต์ใช้แนวคิดลิงค์โอเพนดาทาในการเปิดเผยข้อมูลแบบเสรีบนเครือข่ายอินเทอร์เน็ต

มหาวิทยาลัยเทคโนโลยีสุรนารี