doi: 10.14456/jiskku.2021.15

ระบบสืบค้นเชิงความหมายสำหรับข้อมูลวิจัยทางสารสนเทศศาสตร์

Semantic Search System for Research Data in Information Science

สมเพ็ชร จุลลาบุดดี[1]*, กุลธิดา ท้วมสุข[1]

Sompejch Junlabuddee[1]*, Kulthida Tuamsuk[1]

*Corresponding author email: jsompe@kku.ac.th

Received: March 4, 2021

Revised: June 16, 2021

Accepted: June 18, 2021

บทคัดย่อ

วัตถุประสงค์: การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาออนโทโลยีและระบบสืบค้นเชิงความหมายสำหรับข้อมูลงานวิจัยในสาขาวิชาสารสนเทศศาสตร์ โดยการประยุกต์ใช้ข้อมูลวิจัยจากกลุ่มหัวข้อวิจัยที่ได้จากการวิเคราะห์ด้วยเทคนิคแบบจำลองหัวข้อ

วิธีการศึกษา: ใช้วิธีวิจัยแบบการวิจัยและพัฒนา ประกอบด้วย การวิเคราะห์และจัดกลุ่มเนื้อหา การพัฒนาออนโทโลยีและระบบสืบค้นเชิงความหมาย การประเมินผลโดยผู้เชี่ยวชาญและผู้ใช้ระบบ และการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติ

ข้อค้นพบ: ผลการวิจัยได้ (1) ออนโทโลยีข้อมูลวิจัยในสาขาวิชาสารสนเทศศาสตร์ ประกอบด้วยสามคลาสหลัก คือ ResearchArticle, ResearchTopic มี 30 คลาสย่อย และ ResearchCategory มี 2 คลาสย่อย ประกอบด้วย คุณสมบัติ ความสัมพันธ์ และคำอธิบายรายละเอียด ผลการประเมินคุณภาพของออนโทโลยีโดยผู้เชี่ยวชาญพบว่ามีคุณภาพโดยรวมอยู่ในระดับมาก (2) ระบบสืบค้นเชิงความหมายสำหรับข้อมูลวิจัยทางสารสนเทศ สามารถสืบค้นเชิงความหมายได้ตามคุณสมบัติของคลาส หัวข้อวิจัย หมวดหมู่หัวข้อวิจัย และคำสำคัญ ผลการประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญ พบว่า ในภาพรวมระบบมีประสิทธิภาพอยู่ในระดับมาก ผลการประเมิน

[1] สาขาวิชาสารสนเทศศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น; Department of Information Science, Faculty of Humanities & Social Sciences, Khon Kaen University.

ระบบโดยผู้ใช้ด้วยการวัดตามแนวคิดการค้นคืนความรู้ พบว่า ผลการสืบค้นมีค่าความแม่นยำ ร้อยละ 94.67 ค่าความครบถ้วน ร้อยละ 89.87 และค่าประสิทธิภาพโดยรวม (F-Measure) ร้อยละ 92.21

การประยุกต์ใช้จากการศึกษานี้: (1) ออนโทโลยีข้อมูลวิจัยทางสารสนเทศศาสตร์ ใช้เป็นข้อมูลสำหรับการศึกษาทิศทางและแนวโน้มการทำวิจัยในสาขาวิชา (2) ระบบสืบค้นข้อมูลเชิงความหมายสำหรับข้อมูลวิจัยทางสารสนเทศศาสตร์ สามารถนำไปใช้เป็นต้นแบบในการพัฒนาระบบสำหรับข้อมูลวิจัยในสาขาวิชาอื่น ๆ ได้

คำสำคัญ: ข้อมูลวิจัย บทความวิจัย สารสนเทศศาสตร์ ออนโทโลยี การสืบค้นเชิงความหมาย

Abstract

Purpose of the study: This investigation aimed at developing ontology and semantic search system for research data in information science using the topic modeling method.

Methodology: This research and development study analyzed and classified 30,571 research articles published in internal journals and indexed by Web of Science Database from 2013 to 2019. After the ontology and semantic search system had been developed, they were evaluated by experts and system users, and the assessed data were statistically analyzed.

Main Findings: The ontology comprised 3 main classes: 1) research articles and 2) research topics with 30 subclasses each, and 3) research category which was accompanied by only 2 subclasses. These 3 main classes provided properties, relations and descriptions. The analysis of the assessment of the ontology quality by the experts revealed that its quality and effectiveness as a whole was at the very high level, and the semantic search system for retrieving research data in information science could be conducted through the use of the properties, research topics and category.

Applications of this study: The ontology for research data in information science is of useful direction for determining research trends in information science, and the semantic search system can be used as a guideline for developing research data in other subject fields.

Keyword: Research data, Research article, Information science, Ontology, Semantic search

บทนำ

สารสนเทศศาสตร์ จัดเป็นสหสาขาวิชา (Multidisciplinary) ที่ศึกษาเกี่ยวกับการวิเคราะห์ รวบรวม จัดหมู่ ปรับเปลี่ยน จัดเก็บ ค้นคืน และเผยแพร่สารสนเทศ (Stock & Stock, 2013) ในปี ค.ศ. 1968 The American Documentation Institute ได้เปลี่ยนชื่อมาเป็น American Society for Information Science โดย Borko (1968) ได้นิยามสารสนเทศศาสตร์ว่า เป็นศาสตร์ที่มีพัฒนาการมาจากบรรณารักษศาสตร์ (Library science) โดยเน้นการศึกษาเกี่ยวกับคุณสมบัติและพฤติกรรมสารสนเทศ (Properties and behaviors of information) การใช้และการส่งผ่านสารสนเทศ (Use and transmission of information) และการประมวลผลสารสนเทศ (Information processing) เพื่อให้การเข้าถึงและการใช้สารสนเทศเกิดประโยชน์สูงสุด อย่างไรก็ตาม สารสนเทศศาสตร์ เป็นสาขาวิชาที่มีการเปลี่ยนแปลงตลอดเวลาเนื่องจากได้รับผลกระทบจากความเจริญก้าวหน้าอย่างรวดเร็วของเทคโนโลยีสารสนเทศและการสื่อสาร (Information and communication technology) จึงมีการศึกษาวิจัยเพื่อวิเคราะห์ขอบเขต (Scope) ของสารสนเทศศาสตร์โดยวิธีการที่หลากหลาย เพื่อนำมาใช้เป็นแนวทางในการกำหนดเนื้อหารายวิชาในหลักสูตร การพัฒนาสมรรถนะของนักวิชาชีพสารสนเทศศาสตร์ และการกำหนดประเด็นวิจัยที่สอดคล้องกับการแก้ปัญหาและการพัฒนาที่ใช้สารสนเทศเป็นเครื่องมือสำคัญในการตัดสินใจในมิติต่าง ๆ

การวิจัยเกี่ยวกับสารสนเทศ มีปรากฏในหลักสูตรหลายสาขาวิชา อาทิ สารสนเทศศาสตร์ สารสนเทศศึกษา บรรณารักษศาสตร์ เทคโนโลยีสารสนเทศ การจัดการสารสนเทศ การสื่อสารมวลชน บริหารธุรกิจ เทคโนโลยีการศึกษา หรือแม้แต่ในสาขาวิชาทางวิทยาศาสตร์ เช่น วิศวกรรมศาสตร์ วิทยาศาสตร์สุขภาพ ก็มีการวิจัยเกี่ยวกับสารสนเทศเช่นกัน แต่อาจมีประเด็นการวิจัยในแง่มุมที่แตกต่างกันตามความสัมพันธ์ของประเด็นความสนใจเกี่ยวกับสารสนเทศของแต่ละสาขาวิชา อย่างไรก็ตามสาขาวิชาที่มีจุดเน้นในเรื่องของสารสนเทศ โดยมองว่า สารสนเทศ (Information) เป็นแก่น (Core) ของสาระที่ต้องศึกษาวิจัย โดยมีความเชื่อมโยงกับเทคโนโลยีสมัยใหม่ (Modern technology) ที่นำมาใช้ในการจัดการสารสนเทศและประโยชน์ของการนำไปใช้ในมิติต่าง ๆ ที่ต้องมีผู้ใช้หรือคน (People) เข้ามาเกี่ยวข้องในปัจจุบันเป็นที่เข้าใจตรงกันว่าคือสาขาวิชาสารสนเทศศาสตร์ (Information science) โดยเฉพาะอย่างยิ่งเมื่อมีการตั้ง The iSchools Organization ในปี ค.ศ. 2005 เป็นองค์กรระดับสากลที่มีสมาชิกเป็นสถาบันที่เปิดสอนหลักสูตรสาขาวิชาสารสนเทศศาสตร์หรือที่เกี่ยวข้องจากทั่วโลก ความจำเป็นในการทำวิจัยในสาขาวิชาที่เกี่ยวข้องกับบทบาทที่สำคัญของสารสนเทศที่สอดคล้องกับความต้องในการใช้ชีวิตของประชาชน และการพัฒนาชุมชนและสถาบันยิ่งมีความชัดเจนมากขึ้น (Claver, González & Llopis, 2000; Larsen, 2008; iSchools Inc., 2015)

ข้อมูลวิจัย (Research data) หมายถึง ข้อมูล ข้อเท็จจริงจากการวิจัยที่บันทึกในแหล่งสารสนเทศต่าง ๆ ที่ได้รับการยอมรับในแวดวงวิชาการ และสามารถตรวจสอบผลจากการวิจัยได้ (EPSRC, n.d.) ข้อมูลวิจัยมักขึ้นอยู่กับแนวคิดของแต่ละสาขาวิชา อาทิ งานวิจัยทางมนุษยศาสตร์อาจมีจุดเน้นเรื่องแหล่งข้อมูลปฐมภูมิ

(Primary sources) และอธิบายเนื้อหาเป็นข้อความเชิงพรรณนา (Descriptive texts) งานวิจัยทางสังคมศาสตร์อาจอธิบายเนื้อหาที่เน้นผลลัพธ์จากการศึกษาจากผู้ให้ข้อมูลหรือหลักฐานที่เป็นปรากฏการณ์ทางสังคม โดยวิธีการสำรวจ การสัมภาษณ์ การศึกษาภาคสนาม และการวิเคราะห์ทางสถิติ ส่วนงานวิจัยทางวิทยาศาสตร์อาจเน้นอธิบายขั้นตอนวิธีการ และผลการทดลอง (Monash University Library, 2013) สำหรับการจัดการข้อมูลวิจัย (Research data management) นั้น ก็มีความหลากหลายมาก สถาบันอุดมศึกษาและสถาบันวิจัยต่าง ๆ ซึ่งเป็นแหล่งผลิตผลงานวิจัยโดยตรง มักมีระบบจัดการข้อมูลวิจัยที่อาจจะพัฒนาขึ้นเอง หรือใช้ระบบจัดการข้อมูลแบบเปิด (Open data) ที่เปิดโอกาสให้บุคคลภายนอกเข้าถึงงานวิจัยได้ นอกจากนี้ยังมีระบบจัดการข้อมูลวิจัยของสำนักพิมพ์ต่าง ๆ อีกจำนวนมาก (Beagrie, Lavoie & Woollard, 2009; Sanjeeva, 2018)

สำหรับการค้นคว้างานวิจัยในสาขาวิชาสารสนเทศศาสตร์นั้น หากเป็นบทความวิจัยที่ตีพิมพ์ในวารสารระดับชาติ อาจค้นคว้าได้ไม่ยากนัก เพราะวารสารที่เป็นที่ยอมรับในสาขาวิชาที่และมีชื่อปรากฏในฐานข้อมูล Thai Journal Citation Index (TCI) ในกลุ่ม 1 และกลุ่ม 2 จำนวนไม่มากนัก นอกจากนี้ประเด็นวิจัยก็ยังไม่หลากหลายเท่าที่ควร ดังนั้น การวิจัยในสาขาวิชาสารสนเทศศาสตร์ โดยส่วนใหญ่จะอ้างอิงบทความวิจัยในวารสารระดับนานาชาติ ที่อยู่ในฐานข้อมูล SCOPUS หรือ Web of Science ทั้งนี้อาจเป็นเพราะคุณภาพของงานวิจัยที่จะนำมาใช้อ้างอิงมีความทันสมัยและน่าเชื่อถือมากกว่า นอกจากนี้ สำหรับอาจารย์ในระดับอุดมศึกษา ก็จำเป็นต้องส่งผลงานไปตีพิมพ์ในระดับนานาชาติในฐานข้อมูลทั้งสองแหล่งดังกล่าวข้างต้นตามเกณฑ์ในประกาศ ของคณะกรรมการข้าราชการพลเรือนในสถาบันอุดมศึกษา (Committee for Civil Servants in Higher Education Institutions, 2017, 2020) อีกด้วย ดังนั้น การสร้างสรรค์ผลงานวิจัยในหัวข้อที่มีการตีพิมพ์ในระดับนานาชาติ นับเป็นเรื่องจำเป็น อย่างไรก็ตาม การสืบค้นและการเข้าถึงข้อมูลบทความวิจัยในฐานข้อมูลระดับนานาชาติ ซึ่งมีปริมาณมหาศาล อาจต้องใช้เวลาและอาจไม่ได้ข้อมูลตรงตามความต้องการ

การสืบค้นเชิงความหมาย (Semantic search) เป็นเทคโนโลยีขั้นสูงในการค้นหาความรู้ (Knowledge discovery) โดยใช้เทคโนโลยีเชิงความหมายในการค้นคืนความรู้จากแหล่งข้อมูลที่มีโครงสร้างและไม่มีโครงสร้างด้วยการใช้แบบจำลองกราฟความรู้ (Knowledge graph model) ทำให้ระบบปัญญาประดิษฐ์ (AI system) เข้าใจความหมายของแนวคิด (Concepts) และความสัมพันธ์ระหว่างแนวคิด (Relationship of concepts) ที่ได้มาจากโครงสร้างของออนโทโลยี ทำให้ผลการค้นมีความแม่นยำสูงและมีความเฉพาะเจาะจงสอดคล้องกับสิ่งที่ต้องการค้นได้มากขึ้น (Azati Team, 2020) คำว่า “Semantic” หมายถึง ความหมายหรือสาระของบางอย่าง เมื่อนำมาใช้กับการค้นจึงหมายถึงการศึกษาเกี่ยวกับคำ (ภาษา) และตรรกะของภาษานั้นเพื่อทำความเข้าใจความต้องการของผู้ค้นผ่านความหมายของสาระหรือบริบทของเนื้อหา (Contextual meaning) โดยการจับคู่แนวคิดที่สัมพันธ์กัน (Concept matching) มองหาคำพ้อง (Synonyms) และใช้อัลกอริธึมของภาษาธรรมชาติ (Natural language algorithm) ทำให้แปลงข้อมูลทั้งแบบโครงสร้างและไม่มีโครงสร้างในฐานข้อมูลให้เข้าถึงได้ง่ายและมีความถูกต้องแม่นยำ (Berasategi, 2019; Azati Team, 2020) การสืบค้น

เชิงความหมายขยายขอบเขตจากการสืบค้นเอกสารไปสู่การสืบค้น entity และความรู้ที่อยู่ในเอกสาร ซึ่ง entity หรือ resource สามารถกำหนดเป็น instance ใน class ของออนโทโลยี ที่มีค่าของ attributes และ relations กับ entities อื่น ๆ ได้ ตัวอย่างเช่น ความหมายของคิวรีสำหรับการค้นหา “People in school of information science” จะถูกตีความโดยระบบการสืบค้นเชิงความหมายเป็นการค้นหา บุคคล “Individuals” เช่น professors, lecturers ที่มีความสัมพันธ์กับ School of information science (Sheth, Ramakrishnan, & Thomas, 2005; Wei & Bargiela, 2008) ดังนั้น หัวใจสำคัญหลักของการสืบค้นเชิงความหมาย คือ ออนโทโลยี ซึ่งเป็นเครื่องมือพื้นฐานสำหรับการสืบค้นเชิงความหมาย ความสำเร็จในการสร้างระบบสืบค้นเชิงความหมาย (Semantic search system) ขึ้นอยู่กับจำนวนประโยคที่ถูกนำมาใช้ในการอธิบายความรู้ในขอบเขตนั้น ๆ ให้เข้าใจอย่างชัดเจนและมากเพียงพอด้วยวิธีการออนโทโลยี นักพัฒนาระบบสืบค้นเชิงความหมายสามารถสร้างออนโทโลยีได้ตั้งแต่ขนาดเล็กที่มีความคิดสรุปรวบยอดไว้ หรืออาจเป็นออนโทโลยีขนาดใหญ่ที่มีชุดคำศัพท์และความสัมพันธ์ที่สลับซับซ้อนเชื่อมโยงกันหลายพันคำขึ้นไป ในขอบเขตพื้นที่ที่ข้อมูลมีคุณลักษณะหลากหลายและแตกต่างกัน ดังนั้น เครื่องมือที่จะนำมาสนับสนุนระบบการทำงานในแบบดังกล่าวจึงจำเป็นต้องมีการจัดเตรียมสภาพแวดล้อมให้การทำงานเป็นไปอย่างราบรื่นเพื่อใช้สำหรับการเรียกดู ค้นหา แบ่งปัน และการสร้างข้อมูลสำหรับออนโทโลยีให้สอดคล้องรองรับกับความรู้หรือข้อมูลที่มีอยู่ในเว็บหรือแหล่งข้อมูลในเรื่องนั้น ๆ (Guha, McCool & Miller, 2003; Tuamsuk et al., 2018)

ด้วยเหตุผลและความจำเป็นในการเข้าถึงข้อมูลวิจัยในสาขาวิชาสารสนเทศศาสตร์ดังกล่าวแล้ว ผู้วิจัยจึงมีแนวคิดในการวิจัยเรื่องนี้โดยมีวัตถุประสงค์ เพื่อพัฒนาออนโทโลยีและระบบสืบค้นเชิงความหมายสำหรับการสืบค้นข้อมูลวิจัยในสาขาวิชาสารสนเทศศาสตร์โดยเชื่อมโยงกับรายการบทความที่มีอยู่ในฐานข้อมูลระดับนานาชาติ ด้วยการประยุกต์การทำเหมืองข้อความสำหรับจัดกลุ่มหัวข้อเอกสารงานวิจัย ทำการตรวจสอบและให้นิยามความหมายโดยผู้เชี่ยวชาญ โดยคาดว่าผลจากการวิจัยนี้จะทำให้ได้ (1) ได้ออนโทโลยีสารสนเทศงานวิจัยในสาขาวิชาสารสนเทศศาสตร์ ที่ประกอบด้วยโครงสร้างข้อมูลระดับหมวดหมู่ (Class) และระดับหมวดหมู่ย่อย (Subclass) รวมถึงความสัมพันธ์ของข้อมูลที่มาจากฐานข้อมูลบทความวิจัยระดับนานาชาติ ซึ่งจะเป็นประโยชน์ในการใช้เป็นเครื่องมือสืบค้นข้อมูลเชิงความหมาย และการเข้าถึงความรู้ที่มีในแหล่งข้อมูลต่าง ๆ ทั้งที่มีอยู่บนเว็บและในฐานข้อมูล และ (2) ได้ระบบสืบค้นข้อมูลเชิงความหมายสำหรับการค้นคว้าบทความวิจัยในสาขาวิชาสารสนเทศศาสตร์ ที่มีการจัดเก็บข้อมูลบทความวิจัยตามโครงสร้างรายการบรรณานุกรม ที่ได้มาจากฐานข้อมูลงานวิจัยระดับนานาชาติ และยังสามารถเพิ่มเติมบทความวิจัยจากฐานข้อมูลอื่น ๆ ที่มีการจัดเก็บบทความวิจัยในสาขาวิชาสารสนเทศศาสตร์และสาขาวิชาที่เกี่ยวข้องอีกด้วย

วัตถุประสงค์การวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาออนโทโลยีและระบบสืบค้นเชิงความหมายสำหรับข้อมูลงานวิจัยในสาขาวิชาสารสนเทศศาสตร์ โดยการประยุกต์ใช้ข้อมูลวิจัยจากกลุ่มหัวข้อวิจัยที่ได้จากการวิเคราะห์ด้วยเทคนิคแบบจำลองหัวข้อ

วิธีการศึกษา

การวิจัยนี้ใช้วิธีวิจัยและพัฒนา (Research and development method) ดำเนินการวิจัย 2 ขั้นตอน คือ (1) การพัฒนาออนโทโลยี และ (2) การพัฒนาระบบการสืบค้นเชิงความหมาย สำหรับข้อมูลงานวิจัยทางสารสนเทศศาสตร์ที่นำมาใช้ในงานวิจัยนี้ นำมาจากหัวข้อวิจัยทางสารสนเทศศาสตร์ ที่นำเสนอโดย Junlabuddee & Tuamsuk (2021) ซึ่งได้จากการวิเคราะห์บทความวิจัยที่ตีพิมพ์ในวารสารระดับนานาชาติที่ปรากฏในฐานข้อมูล Web of Science ระหว่างปี ค.ศ. 2013-2019 จำนวน 99 รายชื่อ 677 ฉบับ รวมบทความทั้งสิ้น 30,571 บทความ ซึ่งผลการวิเคราะห์ข้อมูลวิจัยโดยใช้เทคนิคแบบจำลองหัวข้อ (Topic modeling) พบว่าหัวข้อวิจัยทางสารสนเทศศาสตร์ที่พบมีจำนวน 30 หัวข้อ (Topics)

1. การพัฒนาออนโทโลยี

การพัฒนาออนโทโลยี ดำเนินการตามแนวคิดของ (Noy & McGuiness, 2001; Tuamsuk et al., 2018) แบ่งออกได้เป็น 7 ขั้นตอน ดังนี้

1.1 ระบุขอบเขต(Determine scope) เป็นการระบุขอบเขตและวัตถุประสงค์ของการพัฒนาออนโทโลยี โดยงานวิจัยมีเป้าหมายเพื่อการสืบค้นเชิงความหมายข้อมูลวิจัยทางสารสนเทศศาสตร์ เพื่อช่วยให้นักวิจัยหรือผู้สนใจสามารถเข้าถึงข้อมูลงานวิจัยได้หลากหลายมิติ เช่น ชื่อเรื่อง ชื่อวารสาร ปีพิมพ์ ฯลฯ และยังสามารถค้นหาได้ตามกลุ่มหัวข้อวิจัยที่มีการจัดกลุ่มโดยการเรียนรู้ของเครื่องจักร (Machine learning) ด้วยวิธีการสร้างแบบจำลองหัวข้อ ในขั้นตอนนี้ ผู้วิจัยได้ใช้แนวคิดการจัดระบบความรู้ (Knowledge organization) โดยที่หัวข้อที่มีสาระ (Theme) ที่เหมือนหรือใกล้เคียงกันนำมาจัดกลุ่มให้อยู่ภายใต้หมวดใหญ่เดียวกัน แล้วสร้างหมวดใหญ่ที่ประกอบไปด้วยหัวข้อย่อยเหล่านั้น (Tuamsuk et al., 2018) ได้ผลลัพธ์ดังตารางที่ 1

ตารางที่ 1 การจัดกลุ่มหัวข้อวิจัยทางสารสนเทศศาสตร์ที่ได้จากการวิเคราะห์แบบจำลองหัวข้อจากผลงานของ Junlabuddee & Tuamsuk (2021)

ลำดับที่	หัวข้อวิจัย 30 หัวข้อจากงานวิจัยของ Junlabuddee & Tuamsuk (2021)	กลุ่มหัวข้อ 10 กลุ่ม
1	Data analytics (การวิเคราะห์ข้อมูล)	Data Analytic (การวิเคราะห์ข้อมูล)
2	Social media analytics (การวิเคราะห์สื่อสังคม)	

ตารางที่ 1 การจัดกลุ่มหัวข้อวิจัยทางสารสนเทศศาสตร์ที่ได้จากการวิเคราะห์แบบจำลองหัวข้อจากผลงานของ Junlabuddee & Tuamsuk (2021) (ต่อ)

ลำดับที่	หัวข้อวิจัย 30 หัวข้อจากงานวิจัยของ Junlabuddee & Tuamsuk (2021)	กลุ่มหัวข้อ 10 กลุ่ม
3	Bioinformatics (ชีวสารสนเทศศาสตร์)	
4	Data management (การจัดการข้อมูล)	Data Management (การจัดการข้อมูล)
5	Data modeling (การจำลองข้อมูล)	
6	Knowledge transfer (การถ่ายทอดสารสนเทศ)	Knowledge Management (การจัดการความรู้)
7	Knowledge management (การจัดการความรู้)	
8	Knowledge mapping (การทำแผนที่ความรู้)	
9	Knowledge network (เครือข่ายความรู้)	
10	Competency development (การพัฒนาสมรรถนะ)	Information Literacy (การรู้สารสนเทศ)
11	Information literacy (การรู้สารสนเทศ)	
12	Information education (การศึกษาสารสนเทศ)	
13	Knowledge organization (การจัดระบบความรู้)	Digital Content (เนื้อหาดิจิทัล)
14	Digital collection management (การจัดการดิจิทัลคอลเล็กชัน)	
15	Digital resources (ทรัพยากรดิจิทัล)	
16	Digital libraries (ห้องสมุดดิจิทัล)	
17	Information quality (คุณภาพสารสนเทศ)	Information Governance (ธรรมาภิบาลสารสนเทศ)
18	Information security (ความปลอดภัยของสารสนเทศ)	
19	Information governance (ธรรมาภิบาลสารสนเทศ)	
20	Information system (ระบบสารสนเทศ)	System Development (การพัฒนาระบบ)
21	Software development (การพัฒนาซอฟท์แวร์)	
22	Academic library management (การบริหารห้องสมุดอุดมศึกษา)	Library Management (การบริหารห้องสมุด)
23	Scholarly communication (การสื่อสารทางวิชาการ)	Scholarly Communication (การสื่อสารทางวิชาการ)
24	Information development (การพัฒนาสารสนเทศ)	Information Service (การบริการสารสนเทศ)
25	Information economy (เศรษฐศาสตร์สารสนเทศ)	
26	Information retrieval (การค้นคืนสารสนเทศ)	
27	User behavior (พฤติกรรมผู้ใช้)	
28	Information services (บริการสารสนเทศ)	
29	Information marketing (การตลาดสารสนเทศ)	

ตารางที่ 1 การจัดกลุ่มหัวข้อวิจัยทางสารสนเทศศาสตร์ที่ได้จากการวิเคราะห์แบบจำลองหัวข้อจากผลงานของ Junlabuddee & Tuamsuk (2021) (ต่อ)

ลำดับที่	หัวข้อวิจัย 30 หัวข้อจากงานวิจัยของ Junlabuddee & Tuamsuk (2021)	กลุ่มหัวข้อ 10 กลุ่ม
30	Public and Community services (การบริการชุมชนและสาธารณะ)	

1.2 เลือกใช้ตัวแบบที่มีอยู่ (Consider Reuse) งานวิจัยนี้ได้นำแนวคิดเกี่ยวกับกำหนดคำศัพท์ (Term) ของ Doublin Core Metadata (DCMI, 2012) บางส่วนมาใช้ในการกำหนดคำศัพท์ออนโทโลยี เช่น คำศัพท์ที่เกี่ยวกับการกำหนดข้อมูลบรรณานุกรมของบทความวิจัย อาทิ Title, Creator, และ Abstract เป็นต้น

1.3 กำหนดคำศัพท์ (Enumerate Term) เป็นการกำหนดคำศัพท์หรือนิยามสำคัญของออนโทโลยี รวมถึงระบุคุณสมบัติของคำศัพท์แต่ละคำ เช่น Title คือชื่อบทความ, Creator คือชื่อผู้แต่ง, TPTopic คือ ชื่อหัวข้อ เป็นต้น

1.4 กำหนดคลาส (Define Class) เป็นการกำหนดคลาส (Class) ซึ่งประกอบด้วย คลาสหลัก (Super class) และคลาสย่อย (Subclass) รวมทั้งแนวคิด (Concept) ของคลาส เช่น ResearchArticle เป็นคลาสสำหรับจัดเก็บบทความ และ คลาส ResearchTopic เป็นคลาสสำหรับจัดเก็บหัวข้องานวิจัย ResearchCategory เป็นคลาสสำหรับจัดเก็บกลุ่มหัวข้องานวิจัย ดังตัวอย่างตารางที่ 2

ตารางที่ 2 แสดงคลาสหลักและแนวคิดของออนโทโลยีข้อมูลวิจัยทางสารสนเทศศาสตร์

คลาสหลัก (Super Class)	แนวคิด (Concept)
ResearchArticle	บทความวิจัย
ResearchTopic	หัวข้อวิจัย 30 หัวข้อ
ResearchCategory	หมวดหมู่หัวข้อ

1.5 กำหนดคุณสมบัติ (Define Properties) เป็นการกำหนดคุณสมบัติของคลาส และความสัมพันธ์ระหว่างแนวคิดของคลาส ประกอบด้วย (1) ความสัมพันธ์แบบ is-a (จัดเป็น) เช่น Information Retrieval is-a Information Service (2) ความสัมพันธ์ชนิดคุณสมบัติ แบบ part-of (เป็นส่วนของ) เช่น Research topic part-of Research Article และ (3) ความสัมพันธ์แบบ attribute-of (เป็นคุณสมบัติ) เช่น Title attribute-of Research Article ดังตัวอย่างตารางที่ 3

1.6 ระบุข้อกำหนดต่าง ๆ (Define Constraints) เป็นการระบุข้อกำหนดหรือเงื่อนไขให้กับข้อมูล เช่น TPWord มี class constraint เป็นแบบ string หรือ Category มี class constraint แบบ Class: ResearchCategory ดังตัวอย่างตารางที่ 3

ตารางที่ 3 แสดงคุณสมบัติ ความสัมพันธ์ และข้อกำหนดเงื่อนไขของคลาส ResearchArticle

ชื่อคุณสมบัติ (Property name)	ความสัมพันธ์ (Relation)	ข้อกำหนด (Class constrain)	คำอธิบาย (Description)
ArticleID	attribute-of	String	รหัสบทความ
Title	attribute-of	String	ชื่อบทความ
Creator	attribute-of	String	ผู้แต่ง
Abstract	attribute-of	String	บทคัดย่อ
Keyword	attribute-of	String	คำสำคัญ
Source	attribute-of	String	ชื่อวารสาร
Year	attribute-of	String	ปีที่พิมพ์
Identifier	attribute-of	String	รหัสที่ใช้ในการระบุชี้ไปที่เก็บข้อมูล
ResearchTopic	part-of	Class: ResearchTopic	หัวข้องานวิจัย
ResearchCategory	part-of	Class: ResearchCategory	หมวดหมู่หัวข้อวิจัย

1.7 สร้างค่าอินสแตนซ์ของคลาส (Create Instance) เช่น คลาส Information Retrieval มีอินสแตนซ์ เป็นคำสำคัญ ประกอบด้วย search, network, query, web, pattern, structure, node, link, graph, และ keywords.

เมื่อกำหนดรายละเอียดของออนโทโลยีตามกระบวนการข้างต้นเรียบร้อยแล้ว ผู้วิจัยได้ดำเนินการสร้างฐานความรู้ออนโทโลยีด้วยโปรแกรม Hozo-Ontology Editor (http://www.hozo.jp/) ของ University of Osaka ประเทศญี่ปุ่น ที่นำมาใช้ในประเทศไทยโดย ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC, 2012) จากนั้นดำเนินการประเมินออนโทโลยี โดยผู้เชี่ยวชาญที่เป็นนักวิชาการและนักวิจัยด้านออนโทโลยีและเว็บเชิงความหมาย มีความเชี่ยวชาญด้านวิศวกรรมความรู้ ด้านการออกแบบและพัฒนาออนโทโลยี มีผลงานทางวิชาการเป็นที่ยอมรับระดับชาติและนานาชาติ ทั้งหมด 4 ท่าน ผลการประเมินออนโทโลยีมีความสมบูรณ์ถูกต้องเพียงพอที่จะนำไปใช้ในการพัฒนาระบบสืบค้นเชิงความหมายต่อไป

2. การพัฒนาระบบการสืบค้นเชิงความหมาย

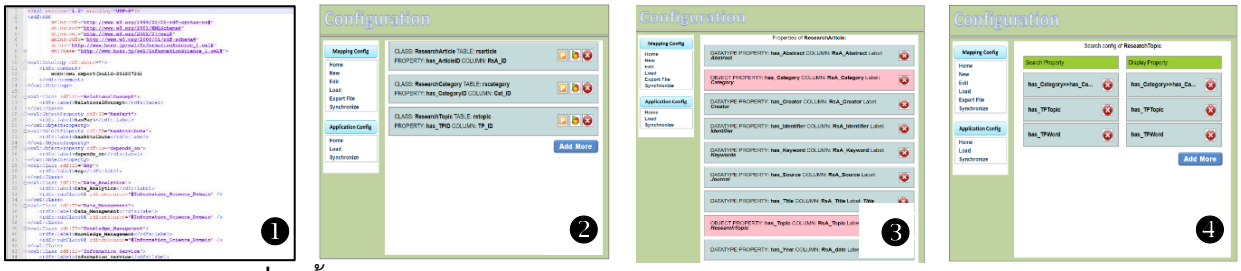
ในขั้นตอนนี้ผู้วิจัยได้นำออนโทโลยีที่ได้จากขั้นตอนการพัฒนาออนโทโลยี มาพัฒนาระบบสืบค้นเว็บเชิงความหมายโดย การประยุกต์ใช้ซอฟต์แวร์แพล็ตฟอร์มสำหรับจัดการโปรแกรมประยุกต์ออนโทโลยีหรือ Ontology Application Management (OAM) Framework พัฒนาโดยห้องปฏิบัติการวิจัยเทคโนโลยี

ภาษาธรรมชาติและความหมาย ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ โดยที่ OAM มีเครื่องมือสนับสนุนการจัดการความรู้เชิงความหมาย (Buranarach et al., 2016) โดยมีขั้นตอนดังนี้

2.1 สร้างเมทาดาตา สำหรับจัดเก็บข้อมูลบทความวิจัยที่สืบค้นได้จากฐานข้อมูล Web of Science โดยเลือกองค์ประกอบ (Elements) ที่จำเป็นจากกรอบมาตรฐาน Dublin Core Metadata Element Set (DCMI, 2012) ประกอบด้วย ชื่อบทความ (Article) ชื่อผู้เขียน (Author) บทคัดย่อ (Abstract) ชื่อวารสาร (Source) ปีพิมพ์ (Date) ที่อยู่ของวารสาร (Identifier)

2.2 ข้อมูล (Instance) ตัวอย่าง เลือกจากข้อมูลบรรณานุกรมของบทความวิจัยที่ได้จัดเก็บไว้แล้วในไฟล์นามสกุล CSV โดยสุ่มบทความจากวารสารทั้ง 99 รายชื่อวารสาร เลือกทุกปีที่มีการตีพิมพ์ในแต่ละวารสารและในปริมาณที่ใกล้เคียงกัน ได้จำนวนบทความทั้งสิ้น 1,020 รายการ

2.3 สร้างระบบสืบค้นข้อมูลเชิงความหมาย โดยใช้โปรแกรม OAM ตามลำดับดังนี้ (1) แปลงฐานความรู้ออนโทโลยีให้อยู่ในรูปภาษา OWL เพื่อนำไปใช้ในการพัฒนาระบบสืบค้นเชิงความหมายเชื่อมโยงออนโทโลยีที่ได้พัฒนาไว้แล้ว กับตัวอย่างชุดข้อมูลบรรณานุกรมของบทความวิจัย (2) กำหนดความสัมพันธ์ของคลาสกับตาราง (Class-table mapping) (3) กำหนดความสัมพันธ์ของคุณสมบัติกับคอลัมน์ (Property-column mapping) และสร้างและจัดเก็บข้อมูลผลลัพธ์ในแบบ RDF (Resource Description Framework) และ (4) ตั้งค่าระบบสืบค้นข้อมูล (Search property) ของคลาสที่อนุญาตให้ระบุอยู่ในเงื่อนไขการสืบค้นได้ และกำหนดค่าการแสดงผล (Display property) ที่จะแสดงอยู่ในผลลัพธ์การสืบค้น ดังภาพที่ 1



ภาพที่ 1 ขั้นตอนการพัฒนาระบบสืบค้นเชิงความหมายโดยใช้โปรแกรม OAM

2.4 ประเมินประสิทธิภาพของระบบสืบค้นข้อมูลเชิงความหมาย การประเมินมุ่งเน้นไปที่ผลลัพธ์ที่ได้จากการแสดงผลของระบบ ด้วยแนวคิด Knowledge retrieval (Yao et al., 2007) โดยการวัดค่าความแม่นยำ (Precision) และค่าความครบถ้วน (Recall) ของการสืบค้น ใช้การประเมินผล โดยภาพรวมสถิติ F-measure (Manning et al., 2008; Vandic et al., 2012) ทำการประเมินโดยผู้ใช้สารสนเทศ 3 ราย และผู้เชี่ยวชาญ 4 ราย ใช้ข้อมูลทดสอบโดยสุ่มเลือกตัวอย่างจากบทความวารสาร ทั้ง 99 วารสาร ได้บทความจำนวน

1,020 บทความ ใช้การประเมินประสิทธิภาพการค้นคืนข้อมูลทั้งค่าความแม่นยำ ค่าความครบถ้วน และประสิทธิภาพในภาพรวมโดยใช้ F-measure และดำเนินการปรับปรุงระบบเป็นขั้นตอนสุดท้าย

ผลการศึกษา

1. ผลการพัฒนาออนโทโลยี ผลการพัฒนาทำให้ได้ “ออนโทโลยีข้อมูลวิจัยทางสารสนเทศศาสตร์ (Ontology of research data in information science)” ซึ่งเกิดจากการนำเอาหัวข้อวิจัยที่ได้จากการวิเคราะห์ข้อมูลบทความวิจัยที่ตีพิมพ์ในวารสารระดับนานาชาติในฐานข้อมูล Web of Science ด้วยเทคนิคแบบจำลองหัวข้อ จำนวน 30 หัวข้อมาพัฒนา ออนโทโลยีมีรายละเอียดดังนี้คือ

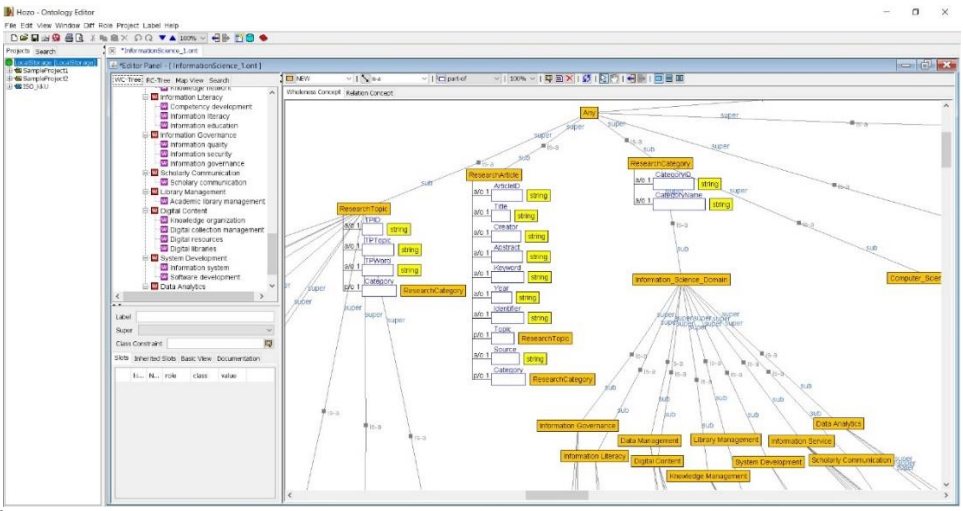
1.1 ออนโทโลยี ประกอบด้วย 3 คลาสหลัก (Super class) ได้แก่ Class: ResearchArticle (บทความวิจัย), Class: ResearchTopic (หัวข้อวิจัย), และ Class: ResearchCategory (หมวดหมู่หัวข้อ)

1.2 คลาสหลัก ประกอบด้วย คลาสย่อย (Subclass) ดังนี้ Class: ResearchArticle ไม่มีคลาสย่อย; Class: ResearchTopic มี 30 คลาสย่อย ได้แก่ หัวข้อวิจัยย่อย 30 หัวข้อ ได้จากการวิเคราะห์แบบจำลองหัวข้อ (ดูตารางที่ 1) และ Class: ResearchCategory มี 2 คลาสย่อย ได้แก่ Information_science_domain และ Computer_science_domain

1.3 คลาสหลัก ประกอบด้วย คุณสมบัติ (Property) ดังนี้ Class: ResearchArticle มีคุณสมบัติได้แก่ ArticleID, Title, Creator, Abstract, Keyword, Source, Year, และ Identifier; Class: ResearchTopic มีคุณสมบัติ ได้แก่ TPID, TPTopic, TPWord, และ Category; และ Class:ResearchCategory มีคุณสมบัติได้แก่ CategoryID และ CategoryName

1.4 ความสัมพันธ์ระหว่างคลาสหลัก คลาสย่อย และคุณสมบัติ มีทั้งแบบ is-a, part_of และ attribute_of

1.5 เมื่อนำมาพัฒนาโดยใช้โปรแกรม Hozo-Ontology Editor ปรากฏโครงสร้างและความสัมพันธ์ ดังภาพที่ 2



ภาพที่ 2 แสดงภาพออนโทโลยีข้อมูลวิจัยทางสารสนเทศศาสตร์ โปรแกรม Hozo-Ontology editor

1.6 ผลการประเมินคุณภาพออนโทโลยีโดยผู้เชี่ยวชาญ สรุปได้ดังนี้ ในภาพรวมออนโทโลยีมีคุณภาพอยู่ในระดับมาก ($\bar{X}$ = 4.19) ผลการประเมินพบว่า การกำหนดคลาสและแนวคิด (Define classes) ($\bar{X}$ = 4.29) การระบุนิยาม ขอบเขตและวัตถุประสงค์ของการพัฒนา (Determine scope) ($\bar{X}$ = 4.25) และการกำหนดคลาสและแนวคิด (Define classes) ($\bar{X}$ = 4.25) มีคุณภาพอยู่ในระดับมากตามลำดับ ส่วนการสร้างอินสแตนซ์ของคลาส (Define instance) ($\bar{X}$ = 4.13) และการกำหนดคุณสมบัติของคลาส (Define properties) ($\bar{X}$ = 4.05) มีคุณภาพอยู่ในระดับมาก (ตารางที่ 4) ซึ่งผู้เชี่ยวชาญมีข้อเสนอแนะเพิ่มเติมคือ (1) การระบุนิยาม ขอบเขตและวัตถุประสงค์ของการพัฒนา ควรมีการกำหนดตามขอบเขตของพื้นที่ (เช่น ประเทศ ภูมิภาคประเทศ) และแหล่งข้อมูลวิจัยว่ามาจากฐานข้อมูลใด (2) การกำหนดคลาสและแนวคิด ควรระบุการอ้างอิงที่มาและ (3) การประยุกต์ใช้และแนวทางการพัฒนาออนโทโลยีในอนาคต ควรสามารถเพิ่มเติมส่วนขยายหรือแนวทางสำหรับใช้ร่วมกันในสาขาวิชาอื่น ๆ และควรเพิ่มข้อมูลวารสารที่มีการเผยแพร่ข้อมูลแบบเปิด (Open access) ได้

ตารางที่ 4 แสดงผลการประเมินคุณภาพออนโทโลยีข้อมูลวิจัยทางสารสนเทศศาสตร์ โดยผู้เชี่ยวชาญ

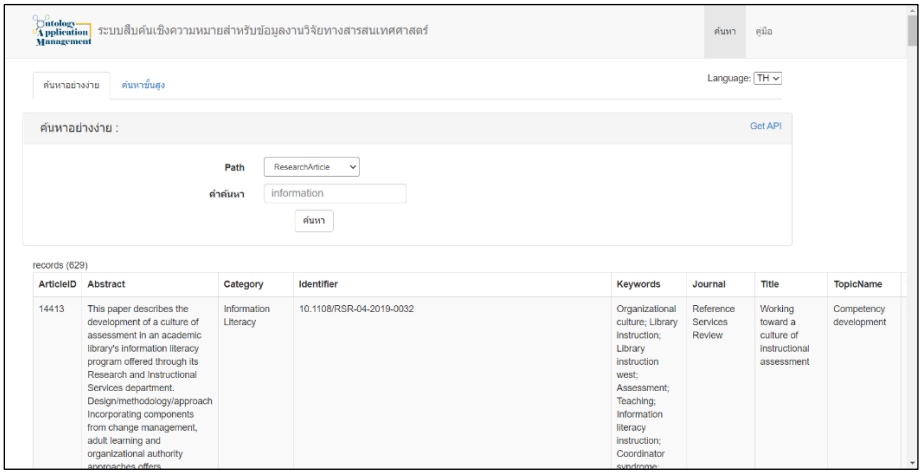
	คุณภาพของออนโทโลยี	ค่าเฉลี่ย	ระดับคุณภาพ
1	การระบุนิยาม ขอบเขตและวัตถุประสงค์ของการพัฒนา (Determine scope)	4.25	มาก
2	การกำหนดคลาสและแนวคิด (Define classes)	4.29	มาก
3	การกำหนดคุณสมบัติของคลาส (Define properties)	4.05	มาก
4	การสร้างอินสแตนซ์ของคลาส (Define instance)	4.13	มาก
5	การประยุกต์ใช้และแนวทางการพัฒนาออนโทโลยีในอนาคต (Future use)	4.25	มาก
	ค่าเฉลี่ยโดยรวม	**4.19**	**มาก**

2. ผลการพัฒนาระบบสืบค้นเชิงความหมาย ผลการพัฒนาทำให้ได้ “ระบบสืบค้นเชิงความหมายสำหรับข้อมูลวิจัยทางสารสนเทศศาสตร์ (Semantic search system for research data in information science)” ซึ่งเกิดจากการนำเอาออนโทโลยีที่พัฒนาขึ้นมาจัดทำระบบโดยใช้โปรแกรม OAM เป็นระบบที่มีคุณลักษณะดังนี้

2.1 สามารถสืบค้นได้ตามคุณสมบัติของคลาส (Search property) และสืบค้นตามหัวข้อ (Topic) 30 หัวข้อที่ได้จากการสร้างแบบจำลองหัวข้อ และการค้นหาตามหมวดหมู่ (Category) 10 หมวดหมู่ของหัวข้อที่ได้จากการจัดกลุ่มใหม่โดยผู้เชี่ยวชาญ

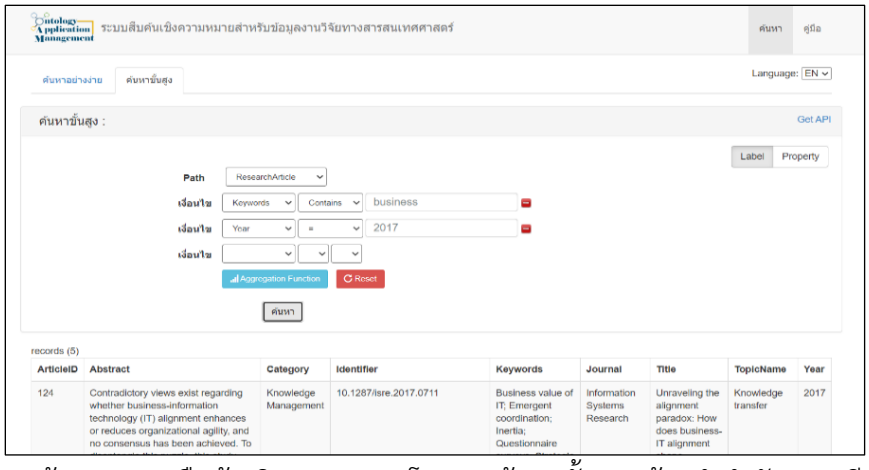
2.2 การใช้งานโปรแกรมระบบสืบค้นข้อมูลเชิงความหมาย สามารถเรียกใช้งานโปรแกรมได้ผ่านโปรแกรมในการท่องเว็บ เช่น Google Chrome จากนั้นพิมพ์ URL http://localhost:8080/searching/

2.3 การสืบค้นอย่างง่าย (Simple search) มีทางเลือกในการสืบค้น (Path) คือ สืบค้นจาก Journal, ResearchArticle, ResearchCategory หรือ ResearchTopic ดังตัวอย่างผลการสืบค้นหาบทความในช่อง Path จาก ResearchArticle และคำค้นที่ใช้ คือ information ปรากฏรายการที่ค้นคืนได้ทั้งหมด 629 รายการ (จาก 1,020 บทความที่ใช้ทดสอบ) พร้อมด้วยรายละเอียดของบทความนั้น ๆ ดังภาพที่ 3

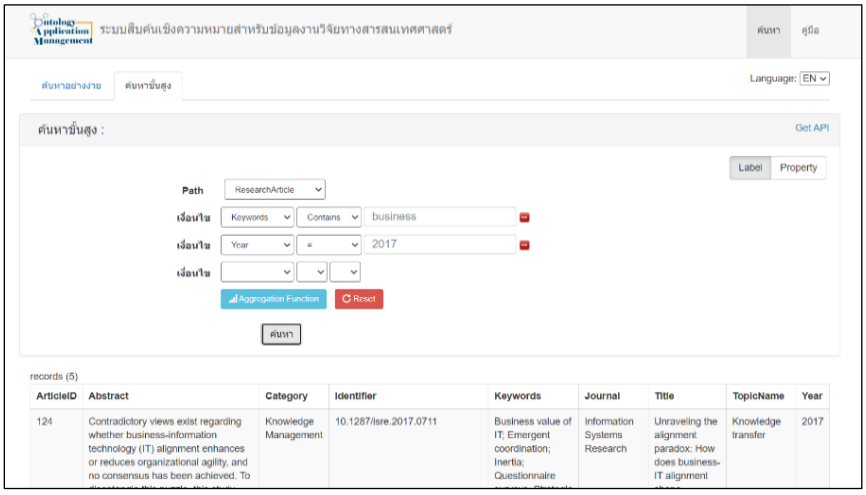


ภาพที่ 3 หน้าแสดงผลการสืบค้นเชิงความหมายโดยการค้นหาอย่างง่าย

2.4 การสืบค้นขั้นสูง (Advanced search) ด้วยการกำหนดเงื่อนไขการสืบค้น สามารถค้นหาข้อมูลโดยเลือกคลาสของสิ่งที่ต้องการค้นหาในช่อง Path ที่ต้องการและเงื่อนไขตามลำดับ ดังตัวอย่างผลการสืบค้นบทความที่มีคำสำคัญ (Keyword) business และพิมพ์ในปี 2017 ปรากฏรายการที่ค้นคืนได้ทั้งหมด 5 รายการ (จาก 1,020 บทความที่ใช้ทดสอบ) พร้อมด้วยรายละเอียดของของบทความนั้น ๆ ดังแสดงในภาพที่ 4 และตัวอย่างการค้นหาบทความที่อยู่ในหัวข้อ (Topic) Data_Analytic และมีคำสำคัญ คือ behavior โดยเลือก Path เป็น ResearchArticle ปรากฏรายการที่ค้นคืนได้ทั้งหมด 3 รายการ (จาก 1,020 บทความที่ใช้ทดสอบ) พร้อมด้วยรายละเอียดของของบทความนั้น ๆ ดังภาพที่ 5



ภาพที่ 4 แสดงหน้าจอผลการสืบค้นเชิงความหมายโดยการค้นหาขั้นสูง ด้วยคำสำคัญ และปีพิมพ์



ภาพที่ 5 แสดงหน้าจอผลการสืบค้นเชิงความหมายโดยการค้นหาขั้นสูง ด้วยหัวข้อ และคำสำคัญ

2.5 การประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญ สรุปผลการประเมินดังนี้ ในภาพรวมระบบสืบค้นเชิงความหมายมีประสิทธิภาพอยู่ในระดับมาก ($\bar{X}$=4.11) ผลการประเมินรายด้านพบว่า ด้านประโยชน์ที่ได้รับ (Perceived usefulness) ($\bar{X}$ = 4.50) และด้านความเหมาะสมแก่เวลา (Timeliness) ($\bar{X}$ = 4.33) มีประสิทธิภาพอยู่ในระดับมาก ส่วนด้านความยากง่ายในการใช้งาน (Ease to use) ($\bar{X}$ = 4.16) ด้านเนื้อหา (Content) ($\bar{X}$ = 4.00) และด้านความถูกต้อง (Accuracy) ($\bar{X}$ = 3.66) มีประสิทธิภาพอยู่ในระดับมาก ตามลำดับ (ตารางที่ 5) ข้อเสนอแนะเพิ่มเติมจากผู้เชี่ยวชาญคือ จุดเชื่อมโยงของ DOI (Digital Object Identifier) หรือ WWW ควรสามารถคลิกและลิงค์ไปยังเนื้อหาของบทความได้เลย และควรมีข้อมูลอันดับควอไทล์ (Quartile) ของวารสารในฐานข้อมูลระดับนานาชาติ และข้อมูลปีที่ ฉบับที่ และเลขหน้าของบทความให้ครบถ้วนตามรายละเอียดข้อมูลบรรณานุกรม

ตารางที่ 5 แสดงผลการประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญ

ประเมินประสิทธิภาพของระบบ		ค่าเฉลี่ย	ระดับประสิทธิภาพ
1.	ด้านเนื้อหา (Content)	4.00	มาก
2.	ความถูกต้อง (Accuracy)	3.66	มาก
3.	ด้านรูปแบบ (Format)	4.00	มาก
4.	ด้านความยากง่ายในการใช้งาน (Ease to use)	4.16	มาก
5.	ด้านความเหมาะสมกับเวลา (Timeliness)	4.33	มาก
6.	ด้านประโยชน์ที่ได้รับ (Perceived usefulness)	4.50	มาก
ค่าเฉลี่ยโดยรวม		**4.11**	**มาก**

2.6 การประเมินประสิทธิภาพโดยผู้ใช้ระบบ โดยมุ่งเน้นไปที่ผลลัพธ์ที่ได้จากการแสดงผลของระบบตามแนวคิด Knowledge retrieval ของ Yao et al. (2007) เพื่อวัดค่าความแม่นยำ และค่าความสมบูรณ์ครบถ้วนของการสืบค้น และคำนวณโดยใช้สถิติ F-measure (Manning et al., 2008; Vandic et al., 2012) ในขั้นตอนนี้ทำการประเมินโดยนักศึกษาระดับบัณฑิตศึกษาและนักวิจัยในสาขาวิชาสารสนเทศศาสตร์ จำนวน 3 คน โดยใช้ข้อมูลทดสอบจากการสุ่มเลือกตัวอย่างบทความจากวารสารทั้ง 99 วารสาร ได้บทความจำนวน 1,020 บทความ กำหนดให้ผู้ประเมินทำการสืบค้นโดยมีเงื่อนไข ดังนี้ (1) ใช้คำสำคัญ (Keyword) ที่ปรากฏภายใต้หัวข้อ (Topic) ที่ถูกจัดหมวดหมู่ไว้แล้ว โดยการเลือกคำสำคัญแบบสุ่ม จำนวนหมวดหมู่ละ 10 คำ จาก 10 หมวดหมู่ รวมคนละ 100 คำ จำนวน 3 คน ทั้งหมด 300 คำ (2) ทำการสืบค้นตามคำสำคัญที่เลือกทีละหมวดหมู่จนครบ 10 หมวดหมู่ (3) ให้ผู้สืบค้นบันทึกว่า คำค้นใดได้ผลลัพธ์ที่ต้องการ คำค้นใดไม่ได้ผลลัพธ์หรือได้ไม่ตรงกับความต้องการ (4) ผลการสืบค้น ได้ผลลัพธ์ของคำที่สืบค้นได้และตรงกับความต้องการ (Relevant) จำนวน 284 คำ และสืบค้นไม่พบหรือไม่ตรงกับความต้องการ (Not relevant) จำนวน 16 คำ นำผลลัพธ์ที่ได้ไปคำนวณดังสมการต่อไปนี้

เมื่อ relevant = จำนวนเอกสารที่เกี่ยวข้องทั้งหมด

retrieved = จำนวนของเอกสารที่เกี่ยวข้องทั้งหมดและถูกดึงออกมา

การหาค่า Precision

$$0.9467 = \frac{284}{300}$$

การหาค่า Recall

$$0.8987 = \frac{284}{16+300}$$

การหาค่า F-measure

$$0.9221 = \frac{2x(0.9467x0.8987)}{0.9467+0.8987}$$

ผลการประเมินได้ค่าระดับความแม่นยำ (Precision) คือ ร้อยละ 94.67 และความครบถ้วน (Recall) คือ ร้อยละ 89.87 ส่วนประสิทธิภาพโดยรวม (F-Measure) เท่ากับ ร้อยละ 92.21 จากผลลัพธ์ที่ปรากกฏให้ค่าความถูกต้องและแม่นยำสูงมาก เท่ากับเป็นการยืนยันประสิทธิภาพของระบบสืบค้นเชิงความหมายที่สามารถนำไปใช้ประโยชน์ได้จริงต่อไป

สรุปผลการศึกษาและอภิปรายผล

การพัฒนาออนโทโลยีในงานวิจัยนี้ ผลลัพธ์ที่ได้ คลาสหลักคือ คลาส ResearchArticle สำหรับจัดเก็บความรู้บทความวิจัยทางสารสนเทศศาสตร์ มีกระบวนการใกล้เคียงกับ Rani, Dhar, & Vyas (2017) คือทำการสกัดคำ สร้างแบบจำลองหัวข้อ (Topic modeling) นำเอาหัวข้อและคำที่ได้จากการสร้างแบบจำลองหัวข้อมาสร้างเป็นออนโทโลยี ส่วนที่แตกต่างกันคือ ในงานวิจัยนี้คือได้คลาส ResearchTopic สำหรับจัดเก็บหัวข้อวิจัย (Topic) และคลาส ResearchCategory สำหรับจัดเก็บหมวดหมู่หัวข้อวิจัย (Category) และให้นิยามความหมายของหัวข้อและกลุ่มหัวข้อโดยอาศัยผู้เชี่ยวชาญในการพิจารณา นอกจากนี้ เมื่อเปรียบเทียบกับออนโทโลยีในสาขาวิชาสารสนเทศศาสตร์ ของ Sawsaa (2013) พบว่า งานของ Sawsaa กำหนดโครงสร้างของออนโทโลยีเฉพาะระดับบนสุดที่เรียกว่า Upper-level หรือ Top-Level ontology ซึ่งมีขอบเขตกว้าง ไม่ได้แบ่งย่อยเนื้อหาในสาขาวิชาสารสนเทศศาสตร์แต่อย่างใด ในขณะที่ออนโทโลยีที่พัฒนาขึ้นในงานวิจัยนี้มีข้อแตกต่างที่สำคัญ คือ (1) เป็นออนโทโลยีข้อมูลวิจัยในสาขาวิชาสารสนเทศศาสตร์ (2) มีการกำหนดคลาสหลักและคลาสย่อยที่แตกหัวข้อวิจัยอย่างเป็นระบบ (3) มีการกำหนดแนวคิด คุณสมบัติ และความสัมพันธ์ของข้อมูลโดยละเอียด และ (4) ใช้โปรแกรมพัฒนาออนโทโลยีซึ่งเป็นที่รู้จักและยอมรับในระดับสากล

การพัฒนาระบบสืบค้นเชิงความหมาย ได้ระบบสืบค้นตามคุณสมบัติของคลาส (Search Property) แนวคิดขั้นตอนการพัฒนาใกล้เคียงกับ Dietze & Schroeder (2009) คือการใช้คีย์เวิร์ดในการสืบค้นและเพิ่มความสามารถด้วยวิธีการทำเหมืองข้อความและออนโทโลยี และทำให้เกิดความง่ายโดยใช้การจัดการด้วยหมวดหมู่ (Category) อย่างไรก็ตาม ในงานวิจัยนี้ได้เพิ่มความสามารถในการค้นหาตามหัวข้อ (Topic) หมวดหมู่หัวข้อ (Category) และคำสำคัญ (Keyword) ที่ได้จากการสร้างแบบจำลองหัวข้อ ซึ่งเป็นการเพิ่มมิติของการค้นหาสารสนเทศงานวิจัยได้และช่วยเพิ่มประสิทธิภาพในการค้นคืนได้มากยิ่งขึ้น

ข้อเสนอแนะเพิ่มเติมจากงานวิจัยนี้คือ (1) ในการพัฒนาออนโทโลยี แม้ผู้วิจัยได้นำองค์ประกอบของ Doublin Core Metadata (DCMI, 2012) มาใช้ในการกำหนดแนวคิดของคลาสและคุณสมบัติแต่ก็ยังไม่สมบูรณ์ครบทุกองค์ประกอบ ดังนั้น เมื่อต้องการให้ออนโทโลยีสามารถเชื่อมโยงกับระบบอื่นที่ใช้ DC ในการอธิบายข้อมูล ก็อาจมีความจำเป็นในการนำองค์ประกอบของ DC ทุกองค์ประกอบมาบรรยายให้ครบ โดยเฉพาะอย่างยิ่งในกรณีที่เป็นข้อมูลบทความวิจัย ซึ่งมีรายละเอียดทางบรรณานุกรมที่เป็นมาตรฐานสากล และ (2) ในการพัฒนาระบบสืบค้นเชิงความหมาย ควรปรับปรุงระบบการสืบค้นที่เมื่อเลือกหมวดหมู่แล้วหัวข้อ

ที่ปรากฏยังแสดงผลไม่ถูกต้องครบถ้วน เช่น เมื่อมีการเลือกหมวดหมู่ของการสืบค้นให้แสดงผลเฉพาะข้อมูลหัวข้อที่ปรากฏในหมวดหมู่นั้นเท่านั้น และเพิ่มจุดเชื่อมโยงข้อมูล DOI หรือ WWW โดยสามารถคลิกและลิงค์ไปยังเนื้อหาของบทความได้ รวมทั้งเพิ่มข้อมูลวารสารที่มีการจัดพิมพ์เผยแพร่แบบเปิดด้วย (3) การนำผลวิจัยนี้ไปใช้สำหรับออนโทโลยีข้อมูลวิจัยทางสารสนเทศศาสตร์ ซึ่งประกอบด้วยโครงสร้างข้อมูลระดับหมวดหมู่ และระดับหมวดหมู่ย่อย รวมถึงความสัมพันธ์ของข้อมูลที่มาจากฐานข้อมูลบทความวิจัยระดับนานาชาติ จะเป็นประโยชน์ในการใช้เป็นข้อมูลสำหรับการศึกษาทิศทางและแนวโน้มการทำวิจัยในสาขาวิชา รวมทั้งนำไปใช้เป็นเครื่องมือสืบค้นข้อมูลเชิงความหมาย และการเข้าถึงความรู้ที่มีในแหล่งข้อมูลต่าง ๆ ทั้งที่มีอยู่บนเว็บและในฐานข้อมูล สำหรับระบบสืบค้นข้อมูลเชิงความหมายสำหรับข้อมูลวิจัยทางสารสนเทศศาสตร์ ที่มีการจัดเก็บข้อมูลบทความวิจัยตามโครงสร้างรายการบรรณานุกรม ที่ได้มาจากฐานข้อมูลงานวิจัยระดับนานาชาติ สามารถนำไปใช้เป็นต้นแบบในการพัฒนาระบบสำหรับข้อมูลวิจัยในสาขาวิชาอื่น ๆ และยังสามารถขยายผลสู่การพัฒนาระบบจัดเก็บและสืบค้นข้อมูลวิจัยที่ปรากฏในวารสารไทยและแหล่งอื่นๆ อาทิ วิทยานิพนธ์ และฐานข้อมูลงานวิจัยของหน่วยงานต่าง ๆ ได้อีกด้วย

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนการวิจัยจากกลุ่มวิจัยมนุษยศาสตร์ดิจิทัล คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น ประจำปีงบประมาณ 2561

เอกสารอ้างอิง

Azati Team. (2020). **What is a semantic search engine and how to build one?.** Retrieved 20 September 2020, from https://www.azati.ai/how-to-build-semantic-search-engine/

Beagrie, N., Lavoie, B., & Woollard, M. (2009). **Keeping search data safe**. Retrieved 2 March 2020, from https://data-archive.ac.uk/media/1687/KRDS2_finalreport.pdf

Berasategi, A. (2019). **Semantic search: What is it? Towards data science**. Retrieved 20 September 2020, from https://www.towardsdatascience.com/semantic-search-73fa1177548f

Borko, H. (1968). Information science: What is it?. **Journal of the Association for Information Science and Technology, 19**(5), 3-5.

Buranarach, M. et al. (2016). OAM: An ontology application management for simplifying ontology-based semantic web application development. **International Journal of Software Engineering and Knowledge Engineering, 26**(1), 115-145.

Claver E., González, R., & Llopis, C. (2000). An analysis of research in information systems (1981-1997). **Information and Management, 37**(4), 181-195.

DCMI. (2012). **Dublin CoreTM metadata element set, version 1.1: Reference description.** Retrieved 5 November 2020, from https://www.dublincore.org/ specifications/dublin-core/dces/

Dietze, H. & Schroeder, M. (2009). **GoWeb: A semantic search engine for the life science web. BMC Bioinformatics**. Retrieved 6 January 2021, from https://doi.org/10.1186/1471-2105-10-S10-S7

EPSRC. (n.d.). **The Engineering and Physical Sciences Research Council (EPSRC) policy framework on research data: Scope and benefits**. Retrieved 16 May 2020, from https://epsrc.ukri.org/about/standards/ researchdata/scope/

Guha, R., McCool, R., & Miller, E. (2003). **Semantic search. Proceedings of the 12th International Conference on World Wide Web**, pp. 700-709. Budapest, Hungary.

iSchools Inc. (2015). **About the iSchools Organization**. Retrieved 21 December 2020, from https://ischools.org/ About.

Junlabuddee, S. & Tuamsuk, K. (2021). Analysis of research data in information science using the topic modeling method. **Journal of Mekong Societies, 17**(1), 89-109.

Larsen, R.L. (2008). **History of the iSchools**. Retrieved 21 December 2020, from https://ischools.org/ resources/Documents/History-of-the-iSchools-2009.pdf

Manning, C.D., Raghavan, P., & Schütze, H. (2009). **An Introduction to Information Retrieval**. Cambridge, UK: Cambridge University Press.

Monash University Library. (2013). **What is research data?.** Retrieved 25 September 2020, from https://www.monash.edu/_data/assets/pdf_file/0010/185869/data-management-brochure.pdf

NECTEC. (2012). **Manual of Hozo-Ontology Editor**. Bangkok: The National Electronics and Computer Technology Center. Retrieved 18 December 2020, from https://lst.nectec.or.th/oam_en/ document_doc/Hozo_ThaiManual_25550123.pdf

Noy, N.F. & McGuinness, D.L. (2001). **Ontology development 101: A guide to creating your first ontology**. Retrieved 12 August 2020, from https://protege.stanford.edu/publications/ontology_development/ ontology101.pdf

Rani, M., Dhar, A.K., & Vyas, O.P. (2017). Semi- automatic terminology ontology learning based on topic modeling. **Engineering Applications of Artificial Intelligence, 63**, 108–125.

Sanjeeva, M. (2018). **Research data management: A new role for academic/research librarians**. Retrieved 2 November 2020, from https://www.researchgate.net/profile/Meghana_Sanjeeva2/publication/323604761_RESEARCH_DATA_MANAGEMENT_A_NEW_ROLE_FOR_ACADEMICRESEARCH_LIBRARIANS/ links/5a9fa7b3a6fdcc22e2cbc5cf/ RESEARCH-DATA-MANAGEMENT-A-NEW-ROLE-FOR-ADEMICRESEARCH- LIBRARIANS.pdf?origin=publication_detail

Sawsaa, A.F. (2013). **Generic model of ontology to visualize information science domain (OIS).** Doctoral Thesis, School of Computing and Engineering, University of Huddersfield. Retrieved 2 August 2020, from http://eprints.hud.ac.uk/id/eprint/17545/

Sheth, A., Ramakrishnan, R., & Thomas, C. (2005). Semantics for the semantic web: The implicit, the formal and the powerful. **International Journal on Semantic Web & Information Systems, 1**(1), 1-18.

Stock, W. G. & Stock, M. (2013). **Handbook of Information Science**. Boston, MA: De Gruyter Saur.

Tuamsuk, K., Chansanam, W., Chaikhambung, J., & Kaewboonma, N. (2018). **Digital Humanities Research.** (In Thai). Khon Kaen: Khon Kaen University.

Vandic, D., van Dam, J.W., & Fasincar, F. (2012). Faceted product search powered by the semantic web. **Decision Support System, 53**(3), 245-437.

Wei, W., Barnaghi, P.M., & Bargiela, A. (2008). **Search with Meanings: An Overview of Semantic Search Systems**. Retrieved 15 August 2020, from http://www.personal.ee.surrey.ac.uk/Personal/ P.Barnaghi/doc/siwn2008Survey.pdf.

Yao, Y., Zeng, Y., Zhong, N., & Huang, X. (2007). Knowledge Retrieval (KR). **IEEE/WIC/ACM International Conference on Web Intelligence (WI'07)**, 729-735, Washington DC, IEEE Computer Society, doi:10.1109/WI.2007.113.