

การคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยพิจารณาถึงคุณภาพการบริการสำหรับการท่องเที่ยว Semantic Web Services Selection by Considering the Quality of Service for tourism

เอกชัย แนนอุดร¹, ศุภกฤษฎี นีวัฒนากุล²

Ekkachai Naenudorn¹, Suphakit Niwattanakul²

Received: 12 July 2017; Accepted: 16 October 2017

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว และเพื่อวิเคราะห์เปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว ปัจจัยที่ใช้ในการพิจารณาถึงคุณภาพการบริการของเว็บเซอร์วิส ได้แก่ เวลาตอบสนอง สภาพความพร้อม ความเสถียรภาพ ค่าบริการ และความสำเร็จในการนำส่งข้อมูล กระบวนการในการดำเนินการวิจัยประกอบไปด้วย การสร้างฐานความรู้ออนโทโลยีของเว็บเซอร์วิสในโดเมนด้านการท่องเที่ยว การจำแนกหมวดหมู่ของเว็บเซอร์วิสโดยใช้หลักการทำให้เหมือน ข้อมูลและเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิส การวัดความคล้ายคลึงกันของเว็บเซอร์วิส การคำนวณหาค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิส การสืบค้นเว็บเซอร์วิสโดยใช้วิธีการถ่วงค่าน้ำหนักของคำ และการค้นหาเว็บเซอร์วิสที่ผู้ใช้สามารถกำหนดเงื่อนไขและค่าน้ำหนักให้กับปัจจัยคุณภาพต่าง ๆ ได้ วิธีการทั้งหมดนี้จะทำให้การคัดเลือกเว็บเซอร์วิสเชิงความหมายมีประสิทธิภาพ และผลลัพธ์ที่ได้จากการค้นหาเว็บเซอร์วิสผู้ใช้จะได้รับเว็บเซอร์วิสที่มีคุณภาพและตรงกับความต้องการของผู้ใช้มากที่สุด

ผลการประเมินผลและวัดประสิทธิภาพโดยใช้วิธีการทำให้เหมือนข้อมูลในการจำแนกหมวดหมู่ของเว็บเซอร์วิส พบว่าเทคนิควิธีที่ให้ค่าความถูกต้อง (Accuracy) สูงที่สุดคือเทคนิควิธีต้นไม้ตัดสินใจ มีค่าเท่ากับ 96.07% รองลงมาคือเทคนิควิธีซัพพอร์ตเวกเตอร์แมชชีน มีค่าเท่ากับ 84.46% เทคนิควิธีเคเนียร์เรสต์เนเบอร์ มีค่าเท่ากับ 84.29% เทคนิควิธีเอนีฟเบย์ มีค่าเท่ากับ 75.18% และเทคนิควิธีเคมีนัส มีค่าเท่ากับ 64.47% ตามลำดับ

ผลการประเมินผลและวัดประสิทธิภาพของระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยว พบว่าผลการประสิทธิภาพโดยการเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้น ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.84 ค่าความระลึก (Recall) เท่ากับ 0.80 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.82 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.93 และผลการประเมินประสิทธิภาพโดยใช้ไดนามิกเกณฑ์ ให้ค่าความแม่นยำเท่ากับ 0.57 ค่าความระลึกเท่ากับ 1.00 ค่าเอฟเมเชอร์เท่ากับ 0.67 และค่าความถูกต้องเท่ากับ 0.93

คำสำคัญ: การคัดเลือกเว็บเซอร์วิส การค้นหาข้อมูลเชิงความหมาย คุณภาพการบริการ การจำแนกหมวดหมู่

Abstract

This research aims to design and develop a QoS-aware semantic web services selection model for tourism as well as to compare the techniques used to classify web services related to tourism services. The factors used to determine service quality of web services are: response time, availability, reliability, cost, and throughput. The research processes include creating an ontology knowledge base for web services in the tourism domain, web services

¹ นักศึกษาปริญญาเอกหลักสูตรวิทยาการสารสนเทศศษญีบัณฑิต สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี

² ผู้ช่วยศาสตราจารย์ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี

¹ Doctoral Student, Doctor of Information Science in Information Technology, Institute of Social Technology, Suranaree University of Technology.

² Assistant professor, Institute of Social Technology, Suranaree University of Technology.

* Corresponding author: e-mail address : ekkachai.n@mbs.msu.ac.th

classification using data mining principles and techniques, measuring the similarity of web services, calculating web services quality scores, using the term weighting method and searching web services where users can define terms and weight values for various quality factors. All of these methods will make the selection of semantic web services effective and the results of web service searches will enable users to get the best quality of the web service that truly meets their needs.

Evaluation results and performance measurements using data mining methods for web services classification found that the best technique was Decision Tree which achieved 96.07% accuracy, followed by 84.46% Support Vector Machine technique, 84.29% k-Nearest Neighbor technique, 75.18% Naive Bayes technique and 64.47% K-means technique, respectively.

The results of the evaluation and measurement of the effectiveness of the search systems for semantic web services according to the quality of service for tourism were that, by comparing between expert and search systems, precision was 0.84, recall was 0.80, F-measure was 0.82, and Accuracy was 0.93. The performance evaluation using dynamic criteria for precision was 0.57, recall was 1.00, F-measure was 0.67 and Accuracy was 0.93.

Keywords: Web services selection, Semantic search, Quality of Service (QoS), Classification.

บทนำ

องค์การการท่องเที่ยวโลกแห่งสหประชาชาติหรือ UNWTO (United Nations World Tourism Organization) ได้คาดการณ์ว่าอนาคตการท่องเที่ยวโลกไปจนถึงปี ค.ศ. 2030 จะเติบโตอย่างต่อเนื่อง โดยตั้งแต่ปี ค.ศ. 1995 เป็นต้นมา จำนวนนักท่องเที่ยวทั่วโลกที่คาดการณ์ไว้ค่อนข้างเป็นไปตามเป้าหมายและในอนาคตจำนวนนักท่องเที่ยวระหว่างประเทศจะเพิ่มขึ้นโดยเฉลี่ยปีละ 43 ล้านคน²⁴ เมื่อการท่องเที่ยวโลกมีแนวโน้มเติบโตมากขึ้นในอนาคต ทำให้มีความต้องการข้อมูลเกี่ยวกับการท่องเที่ยวมากขึ้นตามไปด้วย เช่น ข้อมูลเกี่ยวกับสถานที่ท่องเที่ยว ตำแหน่งที่ตั้ง ที่พัก และข้อมูลเกี่ยวกับระบบคมนาคม เป็นต้น ดังนั้นการนำเสนอข้อมูลด้านการท่องเที่ยวที่มีความถูกต้อง ทันสมัย และรวดเร็วจึงมีความสำคัญ ส่งผลทำให้บุคคลทั่วไปสนใจการท่องเที่ยวมากขึ้นและทำให้เศรษฐกิจด้านการท่องเที่ยวของโลกเติบโตขึ้น

การนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์สามารถแบ่งตามยุคของเว็บได้เป็น 4 ยุค¹⁹ ดังนี้

เว็บในยุคที่ 1 เป็นยุคที่เรียกว่าสแตติกเว็บ (Static Web) ซึ่งเป็นจุดเริ่มต้นของเทคโนโลยีเว็บเกิดขึ้นเมื่อปี ค.ศ. 1994 โดยผู้เริ่มต้นคือ ทิม เบอร์เนอร์ลี (Tim Berners-Lee) ได้ก่อตั้งองค์กรเว็บไซต์สากลขึ้น (World Wide Web Consortium: W3C) เพื่อสร้างมาตรฐานและข้อเสนอแนะสำหรับใช้เป็นหลักในการพัฒนาคุณภาพของเว็บไซต์ การนำเสนอข้อมูลบนเว็บไซต์ในยุคนี้ใช้ภาษาหลักในการพัฒนาเว็บไซต์คือภาษาเอชทีเอ็มแอล (Hyper Text Markup Language: HTML) เป็นการออกแบบเว็บไซต์ที่มีการนำเสนอข้อมูลแบบตายตัวเหมาะกับการนำเสนอข้อมูลที่ไม่มีการเปลี่ยนแปลงบ่อยนัก

การจัดการเนื้อหาทำได้ไม่ค่อยสะดวก

เว็บในยุคที่ 2 เป็นยุคที่เรียกว่าไดนามิกเว็บ (Dynamic Web) ยุคที่สร้างเว็บแอปพลิเคชันเพื่อนำเสนอข้อมูลด้านการท่องเที่ยว ผู้ใช้สามารถโต้ตอบกับเว็บไซต์ได้ เช่น การโพสต์ข้อความ ในยุคนี้ข้อมูลมักถูกจัดเก็บไว้ในฐานข้อมูลทำให้การจัดการข้อมูลทำได้สะดวก มีระบบจัดการข้อมูลหลังร้าน (Back Office) แต่ผู้ดูแลระบบจะต้องเตรียมข้อมูลเองและปรับปรุงข้อมูลให้มีความทันสมัยอยู่ตลอดเวลา การนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ในยุคนี้ใช้ภาษาสคริปต์ในการพัฒนา เช่น PHP, ASP, JSP, Javascript เป็นต้น ส่วนฐานข้อมูลที่นิยมใช้ เช่น Oracle, Microsoft SQL Server, MySQL, PostgreSQL เป็นต้น

เว็บในยุคที่ 3 เป็นยุคที่เรียกว่าเว็บเซอร์วิส (Web Service)²⁸ เป็นการพัฒนาเป็นเว็บแอปพลิเคชันเพื่อเรียกใช้บริการเว็บเซอร์วิสที่ให้บริการข้อมูลด้านการท่องเที่ยว เว็บเซอร์วิสเป็นเทคโนโลยีที่ทำให้โปรแกรมประยุกต์ต่าง ๆ ที่พัฒนาขึ้นโดยมีพื้นฐานทางด้านภาษาและสถาปัตยกรรมที่แตกต่างกัน สามารถสื่อสารแลกเปลี่ยนข้อมูลกันได้โดยใช้ภาษาเอ็กซ์เอ็มแอล (Extensible Markup Language: XML) เป็นภาษากลางในการสื่อสารแลกเปลี่ยนข้อมูลโดยอาศัยโพรโทคอลเอชทีทีพี (Hypertext Transfer Protocol: HTTP) และใช้ร่วมกับมาตรฐานเกี่ยวกับเว็บอื่น ๆ เว็บเซอร์วิสใช้ภาษาดับเบิลยูเอสดีแอล (Web Services Description Language: WSDL) ในการอธิบายถึงรูปแบบการบริการข้อมูล การใช้งานเว็บเซอร์วิสปัจจุบันอาจอยู่ในรูปแบบของโซฟ (Simple Object Access Protocol: SOAP) ตามแบบมาตรฐานแบบดั้งเดิมหรือตามแนวทางแบบเรสท์ (Representational State Transfer:

REST) ที่เรียบง่ายไม่ซับซ้อน การนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ หลายองค์กรเลือกที่จะใช้ข้อมูลด้านการท่องเที่ยวที่ได้รับจากการให้บริการของผู้ให้บริการเว็บเซอร์วิสซึ่งมีความเป็นมืออาชีพและมีความเชี่ยวชาญเฉพาะด้าน ทำให้ได้รับข้อมูลที่ถูกต้องและทันสมัย ซึ่งบางบริการฟรีแต่บางบริการมีการเก็บค่าธรรมเนียมในการใช้บริการ

เว็บในยุคที่ 4 ของการนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ เป็นยุคที่เรียกว่าซีแมนติกเว็บ (Semantic Web) หรือเว็บเชิงความหมาย^{11,26} เป็นพัฒนาการของเวปไซด์ไวต์เว็บที่เกี่ยวข้องกับข้อมูลและการบริการบนเว็บไซต์ ในการสร้างความเป็นไปได้ที่เว็บไซต์จะสามารถเข้าใจถึงความต้องการของผู้ใช้ เว็บเชิงความหมายเป็นวิธีการที่จะทำให้คอมพิวเตอร์หรือแอปพลิเคชันสามารถเข้าใจข้อมูลที่สอดคล้องกับความเข้าใจของมนุษย์ สามารถเข้าใจความหมายของคำและแนวความคิดที่รวบรวมเกี่ยวกับคำในขอบเขตของความรู้ที่ต้องการตามที่ผู้พัฒนาได้กำหนดไว้โดยเชื่อมโยงความสัมพันธ์ของข้อมูลแบบมีโครงสร้าง เว็บเชิงความหมายจึงเป็นอีกแนวคิดหนึ่งที่ถูกนำไปประยุกต์ใช้ในการพัฒนาระบบการค้นคืนสารสนเทศ (Information Retrieval: IR) ในรูปแบบที่เรียกว่าการค้นหาเชิงความหมาย (Semantic Search) ซึ่งเป็นวิธีการค้นหาทรัพยากรสารสนเทศที่ต้องการจากคลังสารสนเทศภายใต้เงื่อนไขหรือกระบวนการคัดเลือกที่ผู้ใช้ต้องการ ทำให้สามารถค้นหาข้อมูลได้ถูกต้องครอบคลุมถึงข้อมูลที่มีความสัมพันธ์เชิงความหมายและได้ข้อมูลที่ตรงตามความต้องการของผู้ใช้มากขึ้น

จากข้อมูลการนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ทั้ง 4 ยุค สรุปได้ว่าเทคโนโลยีเว็บเซอร์วิสและการสืบค้นข้อมูลเชิงความหมายมีความสำคัญต่อการนำไปใช้ในการนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยว เทคโนโลยีเว็บเซอร์วิสทำให้ผู้ใช้สามารถออกแบบและพัฒนาแอปพลิเคชันที่ไม่จำกัดแพลตฟอร์มเพื่อเรียกใช้เว็บเซอร์วิสจากผู้ให้บริการที่เป็นมืออาชีพและมีความเชี่ยวชาญเฉพาะด้าน ทำให้ได้รับข้อมูลที่ถูกต้องและทันสมัย ส่วนการสืบค้นข้อมูลเชิงความหมายจะทำให้ผู้ใช้สามารถค้นหาเว็บเซอร์วิสได้ตรงตามความต้องการของผู้ใช้มากขึ้น

การค้นหาเว็บเซอร์วิสในปัจจุบันผู้ใช้สามารถค้นหาได้จากยูดีดีไอรีจิสตรี (UDDI registry) ฟังก์ชันการทำงานของแต่ละเว็บเซอร์วิสถูกประกาศไว้ในไฟล์ดับเบิลยูเอสดีแอล ซึ่งผู้ให้บริการได้ประกาศไว้บนยูดีดีไอรีจิสตรี ลักษณะของการค้นหาเว็บเซอร์วิสเป็นการใช้คำสำคัญในการค้นหา โดยการนำคำสำคัญไปเปรียบเทียบกับข้อมูลเอกสารเว็บเซอร์วิสที่เก็บในฐานข้อมูล (Keyword Matching) ผลลัพธ์ที่ได้จะพบเฉพาะ

รายการที่มีคำสำคัญตรงกันซึ่งอาจไม่ถูกต้องและครบถ้วน เพราะบางเว็บเซอร์วิสอาจมีความสัมพันธ์กันถึงแม้ว่าเว็บเซอร์วิสนั้นอาจจะไม่มีคำสำคัญนั้นปรากฏในเอกสารเว็บเซอร์วิสเลยก็ตาม ความสัมพันธ์ในลักษณะนี้เรียกว่า ความสัมพันธ์เชิงความหมาย ดังนั้นการสืบค้นเชิงความหมายจึงเป็นแนวทางหนึ่งที่จะช่วยให้การค้นหาข้อมูลมีประสิทธิภาพมากขึ้นเพราะสามารถเข้าใจความหมายที่แท้จริงของเอกสาร

เว็บเซอร์วิสที่ให้บริการข้อมูลด้านการท่องเที่ยวได้รับความสนใจจากผู้ใช้มากขึ้น ดังนั้นจึงมีงานวิจัยที่เกี่ยวข้องกับการคัดเลือกเว็บเซอร์วิสเกิดขึ้นหลายงานแต่ยังมีหลายประเด็นที่ควรปรับปรุงเพื่อให้การคัดเลือกเว็บเซอร์วิสเชิงความหมายที่มีประสิทธิภาพมากขึ้น ผู้วิจัยจึงสรุปประเด็นปัญหาที่สำคัญของงานวิจัยที่เกี่ยวข้องกับการคัดเลือกเว็บเซอร์วิสเชิงความหมายได้ดังนี้

1. งานวิจัยที่เกี่ยวข้องกับการคัดเลือกเว็บเซอร์วิสส่วนใหญ่จะพิจารณาถึงปัจจัยที่มีผลต่อคุณภาพการให้บริการของเว็บเซอร์วิส แต่ไม่ได้มุ่งเน้นการนำเสนอเว็บเซอร์วิสที่ตรงตามความต้องการของผู้ใช้^{2,4,14,32}

2. การคัดเลือกเว็บเซอร์วิสจะพิจารณาคัดเลือกเว็บเซอร์วิสทั้งหมดที่มีอยู่ในฐานข้อมูล ทำให้กระบวนการในการคัดเลือกใช้เวลานาน ซึ่งความเป็นจริงแล้วเว็บเซอร์วิสนั้นมีอยู่เป็นจำนวนมาก ดังนั้นก่อนทำการพิจารณาคัดเลือกด้วยเกณฑ์หรือเทคนิควิธีต่าง ๆ ควรจำแนกหมวดหมู่ของเว็บเซอร์วิสเพื่อให้การคัดเลือกเว็บเซอร์วิสมีประสิทธิภาพมากขึ้น^{5,30,31}

3. ลักษณะของการค้นหาเว็บเซอร์วิสในปัจจุบันจะใช้คำสำคัญจากผู้ไปเปรียบเทียบกับข้อมูลเอกสารเว็บเซอร์วิส (Keyword Matching) ผลลัพธ์ที่ได้อาจไม่ถูกต้องและครบถ้วน^{15,21,22}

4. เว็บเซอร์วิสที่มีให้บริการบนเครือข่ายอินเทอร์เน็ตในปัจจุบันมีจำนวนมากและหลากหลายกลุ่มบริการ จะทราบได้อย่างไรว่าเว็บเซอร์วิสใดในแต่ละกลุ่มบริการมีคุณภาพการบริการดีที่สุด หรือในกรณีที่บางเว็บเซอร์วิสใช้งานไม่ได้ จะมีเว็บเซอร์วิสใดที่สามารถใช้ทดแทนกันได้^{3,7,8}

ดังนั้นผู้วิจัยจึงได้ทำการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการ โดยเน้นไปที่ข้อมูลเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว งานวิจัยนี้ประยุกต์จากข้อมูลกลุ่มการท่องเที่ยวในวารสารขององค์การการท่องเที่ยวโลก²⁵ สามารถแบ่งได้ 5 กลุ่ม ได้แก่ ที่พัก (Accommodation) การเดินทางขนส่ง (Transportation) สถานที่ท่องเที่ยว (Attraction) ตำแหน่งที่ตั้ง (Location) และสภาพอากาศ (Weather) การทำเหมืองข้อมูล

(Data Mining) โดยใช้เทคนิคการจำแนกข้อมูล (Classification) มาใช้ในการจำแนกหมวดหมู่ของเว็บเซอร์วิส คุณภาพการบริการของเว็บเซอร์วิสพิจารณาจาก 5 ปัจจัย ได้แก่ เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ (Response Time) สภาพความพร้อมในการให้บริการ (Availability) ความเสถียรภาพหรือความน่าเชื่อถือ (Reliability) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) และความสำเร็จในการนำส่งข้อมูล (Throughput) งานวิจัยนี้สร้างฐานความรู้โดยใช้ออนโทโลยีของเว็บเซอร์วิสในโดเมนด้านการท่องเที่ยวในการอธิบายความหมายของข้อมูล มีการจำแนกหมวดหมู่ของเว็บเซอร์วิสก่อนที่จะทำการวัดความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่ม จากนั้นวิเคราะห์และเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสเพื่อหาเทคนิควิธีที่มีประสิทธิภาพมากที่สุด มีการคำนวณหาค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิส ขั้นตอนการสืบค้นใช้วิธีการถ่วงค่าน้ำหนักของคำ และในการค้นหาเว็บเซอร์วิสผู้ใช้สามารถกำหนดเงื่อนไขและกำหนดค่าน้ำหนักให้กับปัจจัยคุณภาพต่างๆ ได้อย่างอิสระ (User Preferences) วิธีการทั้งหมดนี้จะทำให้การคัดเลือกเว็บเซอร์วิสเชิงความหมายมีประสิทธิภาพ และผลลัพธ์ที่ได้จากการค้นหาเว็บเซอร์วิสผู้ใช้จะได้รับเว็บเซอร์วิสที่มีคุณภาพและตรงกับความต้องการของผู้ใช้มากที่สุด

วัตถุประสงค์

1. เพื่อออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว
2. เพื่อวิเคราะห์และเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่เกี่ยวข้องกับการบริการด้านการท่องเที่ยว

วิธีดำเนินงานวิจัย

การวิจัยนี้เป็นการศึกษาเชิงประยุกต์ เป็นการศึกษาวิจัยที่มีขั้นตอนการดำเนินงานที่ประยุกต์จากวงจรการพัฒนา ระบบ (System Development Life Cycle: SDLC) มีรายละเอียดดังนี้

1. ศึกษาทฤษฎีและเทคนิคการคัดเลือกเว็บเซอร์วิส ศึกษาทฤษฎีเกี่ยวกับการสืบค้นข้อมูลเชิงความหมาย การจำแนกหมวดหมู่ของเว็บเซอร์วิส การหาความคล้ายคลึงกันของเว็บเซอร์วิส และคุณภาพการให้บริการของเว็บเซอร์วิส
2. การออกแบบสถาปัตยกรรมของระบบ SEMO หรือแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมาย (Semantic web services selection model หรือ SEMO อ่านว่า ซีโม) แสดงได้ดัง Figure 1

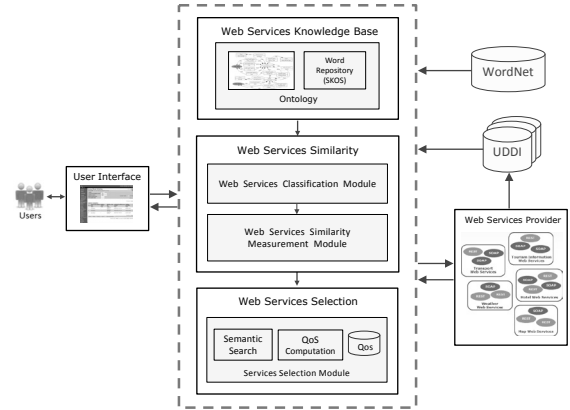


Figure 1 SEMO Architecture

3. การสร้างฐานความรู้โดยใช้ออนโทโลยี

งานวิจัยนี้มีการสร้างฐานความรู้ออนโทโลยีของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว โดยประยุกต์จากออนโทโลยีเว็บเซอร์วิสชื่อ Web Service Modeling Ontology (WSMO)²⁹ ของ W3C แล้วนำมาปรับปรุงเพื่อให้สอดคล้องกับบริบทของการท่องเที่ยว และจัดทำคลังคำโดยใช้สคอส (Simple Knowledge Organization System: SKOS)¹⁹

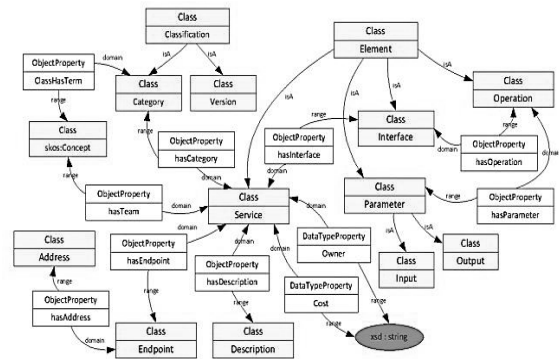


Figure 2 Some of the web services ontology

เก็บรวบรวมคำศัพท์ตัวอย่างที่เกี่ยวข้องกับการบริการด้านการท่องเที่ยวจากพจนานุกรมด้านการสันทนาการ การเดินทางและการท่องเที่ยว (Dictionary of Leisure, Travel and Tourism)¹⁶ จำนวน 200 คำ จากนั้นนำคำศัพท์ตัวอย่างที่ได้มาทำการค้นหาคำศัพท์ที่เกี่ยวข้องจากอภิธานศัพท์เวิร์ดเน็ต (WordNet)¹³ ที่ประกอบไปด้วยคำศัพท์หลัก คำเหมือน คำที่กว้างกว่า คำที่แคบกว่า และคำที่เกี่ยวข้อง และบันทึกคำศัพท์ทั้งหมดที่ได้จากอภิธานศัพท์เวิร์ดเน็ตลงในคลังคำในออนโทโลยีที่ใช้สคอสในโปรแกรมโปรเตจ (Protégé)

4. การจำแนกหมวดหมู่ของเว็บเซอร์วิส (Web Services Classification)

กระบวนการนี้เป็นขั้นตอนการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว ซึ่งกระบวนการในการจำแนกหมวดหมู่ของเว็บเซอร์วิสมี 7 ขั้นตอนดังต่อไปนี้

4.1 เขียนโปรแกรมภาษาพีเอชพี (PHP) ร่วมกับแรพเอพีไอ (RAP API) เพื่อค้นหาข้อมูลเว็บเซอร์วิสจากออนโทโลยีโดยใช้ภาษาสอบถามข้อมูลสปรีย์เคิล (SPARQL)²⁷ ในการสืบค้นข้อมูลในออนโทโลยีที่ถูกส่งออกมาในรูปแบบของภาษาไอบีลยูแอล (OWL) ข้อมูลเว็บเซอร์วิสที่ต้องการได้แก่ ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบายเว็บเซอร์วิส (Description) ชื่อโอเปอเรชัน (Operation) ข้อมูลเข้า (Input) ข้อมูลออก (Output) หมายเลขของเว็บเซอร์วิส (Web Service ID) และคำที่เกี่ยวข้องกับเว็บเซอร์วิส (Term)

4.2 ขั้นตอนการตัดเครื่องหมาย สัญลักษณ์ ตัวเลข และช่องว่าง ออกจากชุดสตริงหลัก หลังจากการสอบถามข้อมูลในขั้นตอนที่ 1 จะได้ข้อมูลเว็บเซอร์วิสที่ต้องการได้แก่ ชื่อเว็บเซอร์วิส คำอธิบายเว็บเซอร์วิส ชื่อโอเปอเรชัน ข้อมูลเข้า ข้อมูลออก หมายเลขของเว็บเซอร์วิส และคำที่เกี่ยวข้องกับเว็บเซอร์วิส ซึ่งข้อมูลทั้งหมดที่ได้จะเป็นชุดสตริงยาว ๆ ที่อาจมีเครื่องหมาย สัญลักษณ์ ตัวเลข หรือช่องว่าง รวมอยู่ด้วย ดังนั้นจึงจำเป็นต้องทำการทำความสะอาดข้อมูล (Data Cleaning) เพื่อกำจัดข้อมูลที่ไม่เกี่ยวข้องออกไป

4.3 การตัดแบ่งคำออกจากสตริง (Word Segmentation) หลังจากการตัดข้อมูลที่ไม่จำเป็นออกในขั้นตอนที่ 2 จะได้ข้อมูลเว็บเซอร์วิสที่สะอาด การตัดแบ่งคำนี้จะรวมถึงการแยกตัวพิมพ์ใหญ่ที่เป็นคำย่อออกด้วย เช่น USA SUT TH SMS API หรือ ISO เป็นต้น ขั้นตอนนี้จะได้ผลลัพธ์ที่เป็นชุดของคำศัพท์ ตัวอย่างเช่น ชุดสตริง "British Airways BA is a full service global airline with an extensive global route network based out of the United Kingdom" เมื่อผ่านขั้นตอนการตัดแบ่งคำนี้ผลลัพธ์ที่ได้จะเป็นคำศัพท์จำนวน 21 คำ ได้แก่ British | Airways | BA | is | a | full | service | global | airline | with | an | extensive | global | route | network | based | out | of | the | United | Kingdom

4.4 การตัดคำหยุด (Stop words) เป็นขั้นตอนตัดคำที่ไม่จำเป็นออก ซึ่งคำหยุดเป็นคำขยายให้แก่คำอื่น ๆ แต่ไม่มีความหมายในตัวเอง เมื่อตัดออกแล้วไม่ทำให้ใจความของเอกสารเปลี่ยนแปลง เช่น คำบุพบทเป็นคำที่ใช้เชื่อมคำหรือกลุ่มคำให้สัมพันธ์กัน คำสรรพนามเป็นคำที่ใช้แทนคำนามที่กล่าวถึงมาแล้วในประโยค คำสันธานเป็นคำที่ทำหน้าที่เชื่อมคำกับคำ เป็นต้น คำหยุดมักเป็นคำที่ปรากฏบ่อยครั้งและมีอยู่ในทุกเว็บเซอร์วิส จึงถือได้ว่าคำหยุดเป็นคุณลักษณะที่ไม่เกี่ยวข้องในการค้นคืนหรือการจำแนกหมวดหมู่ ดังนั้นจึงควร

กำจัดคำหยุดซึ่งจะช่วยประหยัดทั้งพื้นที่และเวลาในการประมวลผล ตัวอย่างคำหยุด เช่น a, an, and, at, but, by, can, do, for, go, give, in, is, of, on, or, other, out, than, the, to, what, while, who, why, will หรือ with เป็นต้น ซึ่งชุดของคำศัพท์ก่อนตัดคำหยุดมีจำนวน 21 คำ หลังจากตัดคำหยุดเหลือเพียง 13 คำ

4.5 การหารากศัพท์ (Stemming Words) ขั้นตอนนี้เป็นกระบวนการในการหารูปเดิมของคำเพื่อปรับรวมให้เป็นคำเดียวกัน การหารากศัพท์ทำให้สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพในการจำแนกหมวดหมู่ของเว็บเซอร์วิสได้ เช่น รากศัพท์ของคำว่า "Booking" คือคำว่า "Book" รากศัพท์ของคำว่า "Transportation" คือคำว่า "Transport" เป็นต้น

4.6 การเปรียบเทียบระหว่างคำศัพท์กับคลังคำในออนโทโลยีเพื่อหาคำอื่นที่เกี่ยวข้อง ขั้นตอนนี้จะนำคำศัพท์แต่ละคำไปเปรียบเทียบกับคลังคำในออนโทโลยีที่จัดเก็บในรูปแบบสคอส (SKOS) เพื่อหาคำหลัก (Prefer Label) คำที่มีความหมายเหมือนกับคำหลัก (Alternate Label) คำที่มีความหมายกว้างกว่า (Broader) และคำที่มีความหมายแคบกว่า (Narrower) จากนั้นนำคำที่ได้ทั้งหมดแทนค่าเพิ่มเข้าไปในชุดสตริงหลัก ตัวอย่างนี้หาคำหลักจำนวน 2 คำ คือ "hotel" และ "reservation" ไปเปรียบเทียบกับคลังคำในออนโทโลยีเพื่อหาคำอื่นที่เกี่ยวข้อง คำหลัก "hotel" ได้คำที่เกี่ยวข้องได้แก่ "building" "edifice" "hostel" "resort" และ "spa" ส่วนคำหลัก "reservation" ได้คำที่เกี่ยวข้องได้แก่ "booking" "employment" "reserve" และ "book"

4.7 เข้าสู่กระบวนการทำเหมืองข้อมูล (Data Mining) เพื่อจำแนกหมวดหมู่ของเว็บเซอร์วิส โดยทำการแบ่งข้อมูลเว็บเซอร์วิสออกเป็น 2 ชุด ดังนี้ ชุดที่ 1 คือชุดข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย ประกอบด้วยแอตทริบิวต์ต่าง ๆ จำนวน 7 แอตทริบิวต์ ได้แก่ รหัสเว็บเซอร์วิส (Web Service ID) ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบาย (Description) ชื่อกลุ่ม (Group Name) ชื่อโอเปอเรชัน (Operation Name) ข้อมูลเข้า (Input) และข้อมูลออก (Output) ส่วนข้อมูลชุดที่ 2 คือชุดข้อมูลของเว็บเซอร์วิสที่ได้ผ่านการหาความสัมพันธ์ของคำเชิงความหมายแล้ว ประกอบด้วยแอตทริบิวต์ต่าง ๆ จำนวน 12 แอตทริบิวต์ ได้แก่ รหัสเว็บเซอร์วิส ชื่อเว็บเซอร์วิส คำอธิบาย ชื่อกลุ่ม (Group Name) ชื่อโอเปอเรชัน ข้อมูลเข้า ข้อมูลออก คำศัพท์หลัก (Terms) คำเหมือน (Alternate Terms) คำที่กว้างกว่า (Broader Terms) คำที่แคบกว่า (Narrower Terms) และคำที่เกี่ยวข้อง (Related Terms) จากนั้นใช้เทคนิควิธีต่าง ๆ

ในการทดลอง เปรียบเทียบด้วย 5 เทคนิควิธี ได้แก่ 1) เทคนิควิธีเคมีนส์ (K-means) 2) เทคนิควิธีเนอ์เบย์ (Naive Bayes : NB) 3) เทคนิควิธีเค-เนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) 4) เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) และ 5) เทคนิควิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine : SVM)

5. การคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส (Web services Similarity)

หลังจากเสร็จขั้นตอนการจำแนกหมวดหมู่ของเว็บเซอร์วิสแล้ว ขั้นตอนนี้จะทำการคำนวณเพื่อหาค่าความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่ม โดยใช้วิธีการหาค่าความคล้ายคลึงตามทฤษฎี Vector Space Model (VSM) คือ การวัดความคล้ายคลึงแบบโคไซน์ (Cosine Similarity)^{17,18} โดยพิจารณาเปรียบเทียบความคล้ายคลึงกันจากชื่อของเว็บเซอร์วิส ชื่อโอเปอเรชัน คำอธิบาย ข้อมูลเข้า และข้อมูลออก แสดงดังสมการที่ 1

$$\text{similarity} = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (1)$$

6. การหาค่าคุณภาพของเว็บเซอร์วิส

วิธีการเลือกเซอร์วิสที่มีความเหมาะสม โดยพิจารณาจาก 5 ปัจจัยดังนี้

1) เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ วัดจากระยะเวลาที่ผู้ขอใช้บริการส่งคำร้องขอออกไปจนกระทั่งได้รับข้อมูลตอบกลับมาจากเซิร์ฟเวอร์ของผู้ให้บริการ ซึ่งคำนวณได้จากสมการที่ 2

$$T = \frac{\sum_{i=1}^n (T_r - T_s)}{n} \quad (2)$$

$$RT = 1 - \frac{T}{\sum T_i}$$

โดยที่ T คือ ค่าเวลาตอบสนองโดยเฉลี่ย หน่วยเป็นมิลลิวินาที

T_r คือ เวลาที่ได้รับข้อมูลกลับมา หน่วยเป็นมิลลิวินาที

T_s คือ เวลาเริ่มต้นที่ทำการส่งคำร้องขอ หน่วยเป็นมิลลิวินาที

RT คือ ค่าความน่าจะเป็นของค่าเวลาตอบสนองโดยเฉลี่ย

2) สภาพความพร้อมในการให้บริการ (Availability) วัดจากอัตราส่วนของผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการสามารถเข้าใช้บริการเซอร์วิสได้สำเร็จส่วนด้วยผลรวมเวลาทั้งหมดที่

ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสในช่วงเวลาที่กำหนด ซึ่งคำนวณได้จากสมการที่ 3

$$A = \frac{N_{suc}}{N_{all}} \quad (3)$$

โดยที่ A คือ ค่าสภาพความพร้อมในการให้บริการ

N_{suc} คือ ผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการสามารถเข้าใช้บริการเซอร์วิสได้สำเร็จ ในช่วงเวลาที่กำหนด มีหน่วยเป็นมิลลิวินาที

N_{all} คือ ผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิส ในช่วงเวลาที่กำหนด มีหน่วยเป็นมิลลิวินาที

3) ความเสถียรของระบบ (Reliability) วัดจากอัตราส่วนของจำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเซอร์วิสได้สำเร็จส่วนด้วยจำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเซอร์วิส ในช่วงเวลาที่กำหนด ซึ่งคำนวณได้จากสมการที่ 4

$$R = \frac{T_{suc}}{T_{all}} \quad (4)$$

โดยที่ R คือ ค่าความเสถียรของระบบ

T_{suc} คือ จำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสได้สำเร็จ ในช่วงเวลาที่กำหนด

T_{all} คือ จำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิส ในช่วงเวลาที่กำหนด

ตัวอย่างเช่น ผู้ใช้บริการร้องขอใช้บริการเซอร์วิสทั้งหมดจำนวน 5 ครั้ง ในช่วงเวลาที่กำหนด และมีสถานการณ์ร้องขอใช้บริการที่ประสบความสำเร็จจำนวน 3 ครั้ง ล้มเหลวจำนวน 2 ครั้ง ดังนั้นค่าความน่าเชื่อถือจะเท่ากับ 3/5 หรือเท่ากับ 0.60

4) ค่าใช้จ่ายในการเรียกใช้บริการ วัดจากอัตราการเก็บค่าบริการในการให้บริการเซอร์วิสของผู้ให้บริการในกลุ่มเดียวกัน ซึ่งคำนวณได้จากสมการที่ 5

$$C = 1 - \frac{P}{\sum P_i} \quad (5)$$

โดยที่ C คือ ค่าความน่าจะเป็นของค่าใช้จ่ายในการเรียกใช้บริการ

P คือ อัตราการเก็บค่าบริการในการให้บริการเซอร์วิส $\sum P_i$ คือ ผลรวมของอัตราการเก็บค่าบริการในการให้บริการเซอร์วิสของผู้ให้บริการทั้งหมด

5) Throughput หมายถึง ความสำเร็จในการนำส่งข้อมูลจากจุดหนึ่งไปยังอีกจุดหนึ่งในช่วงเวลาที่กำหนด มีหน่วยเป็นบิตต่อวินาที (bps) Throughput เป็นจำนวนทรานแซคชัน (Transaction) หรือการร้องขอ (Request) ที่ถูกสร้างขึ้นหรือทำงานได้ในช่วงเวลาการทดสอบหนึ่งๆ ก่อนเริ่มทำการทดสอบ ประสิทธิภาพต้องกำหนดช่วงเวลาไว้ก่อนเสมอ ว่าต้องการเท่าไร เช่น จำนวนใน 1 วินาที หรือจำนวนใน 1 นาที เป็นต้น หน่วยที่นิยมใช้สำหรับวัดค่า Throughput คือ TPS (Transaction Per Second) นั่นคือจำนวนทรานแซคชันที่สามารถทำงานได้ใน 1 วินาที ค่าคะแนนคุณภาพของปัจจัยนี้จะบ่งบอกได้ถึงประสิทธิภาพของการให้บริการเว็บเซอร์วิสโดยเฉพาะในเรื่องขนาดของช่องทางที่ใช้ในการสื่อสารข้อมูลหรือแบนด์วิธ (Bandwidth) เพราะการบริการบางอย่างจำเป็นต้องมีแบนด์วิธเพียงพอจึงจะสามารถใช้งานได้ต้องมีประสิทธิภาพ คำนวณได้จากสมการที่ 6

$$T_p = \frac{N_{suc}}{T_{Total}} \tag{6}$$

$$T_{pmaxTp} = \frac{T_p}{max(T_p)}$$

โดยที่ T_p คือ ค่า Throughput หน่วยเป็น invoke/second

N_{suc} คือ จำนวนทรานแซคชัน (Transaction) หรือการร้องขอ (Request) ที่ถูกสร้างขึ้นหรือทำงานได้ในช่วงเวลาการทดสอบ

T_{total} คือ ค่า ช่วงเวลาการทดสอบ หน่วย second

T_{pn} คือ ค่า Throughput normalize

$max(T_p)$ คือ ค่า Throughput ที่มีค่าสูงสุด

จากนั้นทำการคำนวณค่าคุณภาพการให้บริการของเว็บเซอร์วิส โดยใช้สมการถ่วงน้ำหนัก คำนวณจากการรวมค่าคุณภาพการให้บริการของทั้ง 5 ปัจจัย สูตรการคำนวณดังสมการที่ 7

$$QoS_{si} = \sum_{x \in V} W_x QoS_{si,x} \tag{7}$$

โดยที่ QoS_{si} คือ ค่าคุณภาพการให้บริการของเว็บเซอร์วิส $QoS_{si,x}$ คือ ค่าคุณภาพการให้บริการด้านปัจจัย x ของเว็บเซอร์วิส S_i

W_x คือ ค่าน้ำหนักที่ผู้ใช้บริการกำหนดให้กับคุณภาพด้านปัจจัย x

V คือ เซตของปัจจัยที่ใช้กำหนดคุณภาพการ

ให้บริการของเซอร์วิส โดยที่ $V = \{Response\ Time, Availability, Reliability, Cost, Throughput\}$

ตัวอย่างการคำนวณค่าคุณภาพการให้บริการของเว็บเซอร์วิส โดยมีค่าเริ่มต้นของน้ำหนักปัจจัยดังนี้ Response Time เท่ากับ 1.00 Reliability เท่ากับ 0.75 Availability เท่ากับ 0.75 Cost เท่ากับ 0.50 และ Throughput เท่ากับ 0.75 ซึ่งค่าน้ำหนักของทั้ง 5 ปัจจัยนี้ผู้ใช้สามารถกำหนดค่าได้เองตามความเหมาะสมได้ที่เมนู "Setting" บนหน้าเว็บหลัก

7) การออกแบบและพัฒนาระบบการค้นหาเว็บเซอร์วิสเชิงความหมาย

กระบวนการในการการออกแบบและพัฒนากระบวนการค้นหาเว็บเซอร์วิสเชิงความหมาย มีขั้นตอนดังต่อไปนี้

1. ออกแบบและพัฒนาส่วนติดต่อกับผู้ใช้ (User Interface) ในการค้นหาเว็บเซอร์วิสเชิงความหมาย

ในการค้นหาข้อมูลเว็บเซอร์วิส ผู้ใช้ต้องใส่คำสำคัญ (Keyword) เพื่อสืบค้น จากนั้นระบบจะทำการเปรียบเทียบระหว่างคำสำคัญกับฐานความรู้ออนโทโลยีของเว็บเซอร์วิส ดังนั้นจึงต้องมีการพัฒนาส่วนติดต่อกับผู้ใช้ เพื่อเป็นตัวกลางในการติดต่อกันระหว่างผู้ใช้และฐานความรู้ออนโทโลยี โดยใช้ภาษาพีเอชพี (PHP) เป็นภาษาหลักในการพัฒนาส่วนติดต่อกับผู้ใช้ ใช้ซีเอสเอส (Cascading Style Sheet : CSS) ในการจัดรูปแบบการแสดงผล ส่วนการสืบค้นข้อมูลจากออนโทโลยีใช้ภาษาสปรินกิล เพื่อสอบถามข้อมูล และดึงข้อมูลจากโอดับเบิลยูแอล ผ่านทางแรพเอพีไอ

2. การกำหนดค่าน้ำหนักของคำในโครงสร้างสคอส (SKOS)

ในการกำหนดค่าน้ำหนักที่ถูกจัดเก็บตามโครงสร้างแบบสคอส ซึ่งค่าน้ำหนักนี้เป็นการกำหนดค่าเริ่มต้นทั้งนี้ผู้ใช้สามารถกำหนดเองได้ภายหลัง ซึ่งสามารถกำหนดได้ในเมนู "Setting" บนหน้าเว็บหลัก

3. การกำหนดค่าน้ำหนักของคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส

ในการกำหนดค่าน้ำหนักของคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส ซึ่งคำเหล่านี้ถูกจัดเก็บไว้ในรูปแบบของคลาสในออนโทโลยีแล้ว ได้แก่ ชื่อเว็บเซอร์วิส คำอธิบายเว็บเซอร์วิส ชื่อโอเปอเรชัน ข้อมูลเข้า และข้อมูลออก ทั้งนี้ผู้ใช้สามารถกำหนดค่าน้ำหนักได้เองตามความเหมาะสมในภายหลัง ซึ่งสามารถกำหนดได้ในเมนู "Setting" บนหน้าเว็บหลัก

4. การกำหนดค่าน้ำหนักของคำที่ใช้สำหรับสืบค้น

ในการกำหนดค่าน้ำหนักของคำที่ใช้สำหรับสืบค้นจะพิจารณาจากลำดับของคำสำคัญที่ผู้ใช้ป้อนเข้ามาจากซ้ายไปขวา เช่น ผู้ใช้ป้อนคำว่า “hostel booking” คำว่า hostel จะมีความสำคัญและมีค่าน้ำหนักมากกว่าคำว่า booking เป็นต้น ในการทดสอบนี้จะกำหนดให้คำสำคัญลำดับแรกมีค่าน้ำหนักคือ 1.00 และคำสำคัญลำดับที่สองมีค่าน้ำหนักคือ 0.75 ค่าน้ำหนักนี้สามารถปรับเปลี่ยนได้ตามความเหมาะสมในกรณีที่คำสำคัญมีมากกว่า 2 คำขึ้นไป

5. คำนวณหาค่าน้ำหนักของคำทุกคำที่อยู่ในกลุ่มคำที่ได้จากข้อ 3 เมื่อเปรียบเทียบกับค่าน้ำหนักที่ถูกจัดเก็บตามโครงสร้างแบบสคอส สามารถแทนค่าได้ตามสมการที่ 8

$$SKOSPOS = \sum_{k=0}^n (WSP_k * WePoSK_k) \quad (8)$$

โดยที่ SKOSPOS คือ ค่าน้ำหนักของคำที่อยู่ในโครงสร้างของสคอส

WSP คือ ค่าน้ำหนักของคำที่อยู่ในแต่ละตำแหน่งตามโครงสร้างของเว็บเซอร์วิส

WePoSK (Weight of Position SKOS) คือ ค่าน้ำหนักประจำตำแหน่งในโครงสร้างของเว็บเซอร์วิส

6. การคำนวณหาค่าความคล้ายคลึงกันเชิงความหมายระหว่างคำสำคัญกับออนโทโลยีเว็บเซอร์วิส สามารถแทนค่าได้ตามสมการดังนี้

$$SIM = \sum_{k=0}^n (SKOSPOS_k * WPQ_k) \quad (9)$$

โดยที่ SIM คือ ค่าคะแนนเชิงความหมายของแต่ละเว็บเซอร์วิส SKOSPOS คือ ค่าน้ำหนักของคำที่อยู่ในโครงสร้างของ SKOS

WPQ (Weight of Position Query) คือ ค่าน้ำหนักประจำตำแหน่งของคำสำคัญที่ได้จากผู้ใช้

$$SEMSIM = \frac{SIM_k}{\max(SIM)} \quad (10)$$

โดยที่ SEMSIM คือ ค่าคะแนนความคล้ายคลึงกันเชิงความหมายหลังจากผ่านการปรับปรุงค่า (Normalization) โดยค่าคะแนนจะอยู่ระหว่าง 0 ถึง 1

SIM คือ ค่าคะแนนเชิงความหมายของแต่ละเว็บเซอร์วิส

max(SIM) คือ ค่าสูงสุดของค่าคะแนนเชิงความหมายของเว็บเซอร์วิสทั้งหมด

7. ขั้นตอนการค้นหาเว็บเซอร์วิสเชิงความหมาย ผู้วิจัยทำการออกแบบขั้นตอนการค้นหาเว็บเซอร์วิสเชิงความหมาย

$$QSS = \left[\frac{(Q * W_Q) + (S * W_S)}{W_Q + W_S} \right] \quad (11)$$

โดยที่ QSS คือ ค่าผลรวมคะแนนคุณภาพ (QoS) และค่าความคล้ายคลึง (Similarity) แบบถ่วงน้ำหนัก

Q คือ ค่าคะแนนคุณภาพ (QoS)

S คือ ค่าคะแนนความคล้ายคลึง (Similarity)

W_Q คือ ค่าน้ำหนักสำหรับค่าคุณภาพ (QoS)

W_S คือ ค่าน้ำหนักสำหรับค่าความคล้ายคลึง (Similarity)

$$RankSum(W_i) = \frac{n - r_j + 1}{\sum(n - r_k + 1)} \quad (12)$$

โดยที่ W_i คือ ค่าน้ำหนักของข้อมูล

r_j คือ ลำดับความสำคัญของข้อมูล

r_k คือ ผลรวมของค่าของข้อมูล

n คือ จำนวนข้อมูลทั้งหมด

8. การประเมินผล

8.1 การประเมินผลการสืบค้นเว็บเซอร์วิสเชิงความหมาย

เมื่อทำการสืบค้นเว็บเซอร์วิสจากออนโทโลยี ระบบจะแบ่งกลุ่มของเว็บเซอร์วิสออกเป็น 2 ส่วน คือ เว็บเซอร์วิสที่ถูกดึงออกมาเป็นผลลัพธ์จากการสืบค้น (Retrieved) และเว็บเซอร์วิสที่ไม่ถูกดึงออกมา (Not Retrieved) ซึ่งเว็บเซอร์วิสต่าง ๆ ใน 2 กลุ่มนี้ อาจมีทั้งเว็บเซอร์วิสที่เกี่ยวข้องมีคุณสมบัติผ่านเกณฑ์การเปรียบเทียบความใกล้เคียงของการสืบค้นครั้งนั้น หรือคือเว็บเซอร์วิสที่ผู้ใช้ต้องการ (Relevant) และไม่เกี่ยวข้อง (Irrelevant) กับสิ่งที่ต้องการ ซึ่งสามารถแบ่งกลุ่มของเว็บเซอร์วิสทั้งหมดออกเป็น 4 กลุ่ม คือ

1) Retrieved and Relevant คือกลุ่มเว็บเซอร์วิสที่ถูกดึงออกมาเป็นผลลัพธ์และเกี่ยวข้องกับสิ่งที่ต้องการ (True Positive: TP)

2) Not Retrieved and Irrelevant คือกลุ่มเว็บเซอร์วิสที่ไม่ถูกดึงออกมาเป็นผลลัพธ์ และไม่เกี่ยวข้องกับ

สิ่งที่ต้องการ (True Negative: TN)

3) Retrieved and Irrelevant คือกลุ่มเว็บเซอร์วิสที่ถูกดึงมาเป็นผลลัพธ์ แต่ไม่เกี่ยวข้องกับสิ่งที่ต้องการ (False Positive: FP)

4) Not Retrieved and Relevant คือกลุ่มเว็บเซอร์วิสที่ไม่ถูกดึงมาเป็นผลลัพธ์ แต่เกี่ยวข้องกับสิ่งที่ต้องการ (False Negative: FN)

จากการสืบค้นเว็บเซอร์วิสจะเห็นได้ว่าการสืบค้นนั้นไม่สามารถค้นคืนเว็บเซอร์วิสที่เกี่ยวข้องกับสิ่งที่ผู้ใช้งานต้องการได้ 100% จึงมีการคิดค่าการประเมินผลการสืบค้นเว็บเซอร์วิสได้แก่ ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าเอฟเมเชอร์ (F-Measure)¹²

ค่าความแม่นยำ (Precision) คือ การพิจารณาความถูกต้องของข้อมูลที่สืบค้นได้ จากอัตราส่วนระหว่างจำนวนเว็บเซอร์วิสที่ค้นคืนถูกต้องกับจำนวนเว็บเซอร์วิสที่ค้นคืนมาได้ทั้งหมด (ตัวอย่าง: เมื่อค่าความแม่นยำเท่ากับ 1 แสดงว่าเว็บเซอร์วิสที่ค้นคืนมาได้ทั้งหมดเป็นผลลัพธ์ที่ต้องการ) สูตรการหาค่าความแม่นยำแสดงดังสมการที่ 13

$$\text{Precision} = \frac{tp}{tp + fp} \quad (13)$$

โดยที่ Precision คือ ค่าความแม่นยำในการสืบค้นเว็บเซอร์วิส
 tp คือ เว็บเซอร์วิสที่เกี่ยวข้องและถูกสืบค้นได้
 fp คือ เว็บเซอร์วิสที่ไม่เกี่ยวข้องแต่สืบค้นได้

ค่าความระลึก (Recall) คือ การพิจารณาความครบถ้วนของข้อมูลเมื่อเทียบกับข้อมูลที่ควรได้ทั้งหมด จากอัตราส่วนระหว่างจำนวนเว็บเซอร์วิสที่ค้นคืนถูกต้องกับจำนวนเว็บเซอร์วิสที่ต้องการทั้งหมดในระบบ (หมายเหตุ: ค่าความระลึกเน้นให้คำตอบที่ถูกต้องออกมาให้ได้มากที่สุด) สูตรการหาแสดงดังสมการที่ 14

$$\text{Recall} = \frac{tp}{tp + fn} \quad (14)$$

โดยที่ Recall คือ ค่าความระลึกในการสืบค้นเว็บเซอร์วิส
 tp คือ เว็บเซอร์วิสที่เกี่ยวข้องและถูกสืบค้นได้
 fn คือ เว็บเซอร์วิสที่เกี่ยวข้องแต่สืบค้นไม่ได้

นอกจากการทดสอบประสิทธิภาพของระบบเพื่อวัดความถูกต้องและแม่นยำด้วยค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) แล้ว ยังมีการวัดคุณภาพของผลการค้นหาข้อมูลโดยรวมด้วยค่าเอฟเมเชอร์ (F-Measure) ซึ่งเป็นค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก แสดงดังสมการที่ 15

$$\text{F-measure} = 2 \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (15)$$

โดยที่ F-measure คือ ค่าเอฟเมเชอร์ที่วัดคุณภาพของผลการค้นหาโดยรวม

precision คือ ค่าความแม่นยำในการสืบค้นเว็บเซอร์วิส

recall คือ ค่าความระลึกในการสืบค้นเว็บเซอร์วิส

8.2 การประเมินผลการจำแนกหมวดหมู่ของเว็บเซอร์วิส

โดยใช้การทำเหมืองข้อมูล (Data Mining) และใช้เทคนิควิธีทั้ง 5 เทคนิคนี้ ได้แก่

1) เทคนิควิธีเคมีนส์ (K-means) โดยกำหนดค่า k เท่ากับ 5

2) เทคนิควิธีเนอว์เบย์ (Naive Bayes : NB) โดยกำหนดค่า cross-validation เท่ากับ 10

3) เทคนิควิธีเค-เนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) โดยกำหนดค่า k เท่ากับ 5 และกำหนดค่า cross-validation เท่ากับ 10

4) เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) โดยกำหนดค่า cross-validation เท่ากับ 10

5) เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) โดยกำหนดค่า cross-validation เท่ากับ 10

จากนั้นประเมินผลการจำแนกหมวดหมู่ของเว็บเซอร์วิส โดยการคำนวณหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ตามสมการที่ 13 สมการที่ 14 สมการที่ 15 และตามสมการที่ 16

$$\text{Accuracy} = \frac{tp+tn}{tp+tn+fp+fn} \quad (16)$$

โดยที่ Accuracy คือ ค่าความถูกต้องในการสืบค้นเว็บเซอร์วิส
 tp คือ เว็บเซอร์วิสที่เกี่ยวข้องและถูกสืบค้นได้
 tn คือ เว็บเซอร์วิสที่ไม่เกี่ยวข้องและสืบค้นไม่ได้
 fp คือ เว็บเซอร์วิสที่ไม่เกี่ยวข้องแต่สืบค้นได้
 fn คือ เว็บเซอร์วิสที่เกี่ยวข้องแต่สืบค้นไม่ได้

ผลการวิจัย

1. ผลการจำแนกหมวดหมู่ของเว็บเซอร์วิสและการประเมินประสิทธิภาพ

เพื่อเป็นการแสดงให้เห็นถึงประสิทธิภาพของ

การใช้ฐานความรู้แบบออนโทโลยีในการจัดการข้อมูลเว็บเซอร์วิสเชิงความหมาย โดยผู้วิจัยได้ทำการเก็บรวบรวมเว็บเซอร์วิสจาก Public UDDI Registry เช่น www.webserVICEX.net , www.programmableweb.com และ www.odata.org เป็นต้น จากจำนวนทั้งหมด 200 เว็บเซอร์วิส ทำการคัดเลือกเฉพาะเว็บเซอร์วิสที่มีความสมบูรณ์ที่สุดจำนวน 76 เว็บเซอร์วิส เพื่อนำมาใช้ในการจำแนกหมวดหมู่ของเว็บเซอร์วิส โดยการคัดเลือกเว็บเซอร์วิสที่มีความสมบูรณ์พิจารณาจากเว็บเซอร์วิสที่มีองค์ประกอบครบถ้วน เช่น คำอธิบายเว็บเซอร์วิส โอเปอเรชัน ข้อมูลเข้า ข้อมูลออก และเว็บเซอร์วิสนั้นยังอยู่ในสถานะที่ใช้งานได้ (Active) จากนั้นนำข้อมูลเว็บเซอร์วิสที่ได้ผ่านกระบวนการต่าง ๆ เช่น การตัดแบ่งคำ (Word Segmentation) การตัดคำหยุด (Stop words) และการหารากศัพท์ (Stemming Words) เป็นต้น โดยแบ่งข้อมูลของเว็บเซอร์วิส ออกเป็น 2 รูปแบบ คือ ข้อมูลในชุดที่ 1 คือข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย และข้อมูลในชุดที่ 2 คือ ข้อมูลของเว็บเซอร์วิสที่ได้ผ่านการหาความสัมพันธ์ของคำในออนโทโลยี จากนั้นทำการเปรียบเทียบการจำแนกข้อมูลเว็บเซอร์วิสระหว่างข้อมูลทั้งสองชุด ผลการทดลองสามารถแสดงดังนี้

จากผลการทดลองการจับกลุ่มเว็บเซอร์วิสพบว่า ในชุดข้อมูลแบบที่ 1 เทคนิควิธีที่ให้ค่าความถูกต้อง (Accuracy) สูงที่สุดคือเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) คือ 82.86% รองลงมาคือเทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) คือ 81.61% เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) คือ 75.00% เทคนิควิธีเนออีฟเบย์ (Naive Bayes : NB) คือ 69.46% และเทคนิควิธีเคมีนส์ (K-means) คือ 64.47%

ในชุดข้อมูลแบบที่ 2 เทคนิควิธีที่ให้ค่าความถูกต้อง (Accuracy) สูงที่สุดคือเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) คือ 96.07% รองลงมาคือเทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) คือ 84.46% เทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) คือ 84.29% เทคนิควิธีเนออีฟเบย์ (Naive Bayes : NB) คือ 75.18% และเทคนิควิธีเคมีนส์ (K-means) คือ 64.47%

เทคนิควิธีต้นไม้ตัดสินใจให้ค่าความถูกต้องสูงที่สุด และตามด้วยเทคนิควิธีเคเนียร์เรสท์เนเบอร์และซัพพอร์ทเวกเตอร์แมชชีนที่ให้ค่าความถูกต้องรองลงมา เนื่องจากเป็นเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) ซึ่งเป็นประเภทที่รู้เป้าหมายล่วงหน้าและเหมาะกับข้อมูลเว็บเซอร์วิสที่มีแอตทริบิวต์จำนวนมาก ส่วนเทคนิควิธีเคมีนส์ที่ให้ค่าความถูกต้องน้อยที่สุด ซึ่งจัดอยู่ในเทคนิคแบบการ

เรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) ทำให้ไม่เหมาะกับข้อมูลที่มีแอตทริบิวต์จำนวนมาก

จากการเปรียบเทียบค่าความถูกต้อง (Accuracy) ของชุดข้อมูลทั้งสองชุด พบว่าชุดข้อมูลในแบบที่ 2 (ข้อมูลของเว็บเซอร์วิสที่ได้ผ่านการหาความสัมพันธ์ของคำในออนโทโลยี) ให้ค่าความถูกต้องที่สูงกว่าชุดข้อมูลในแบบที่ 1 (ข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำ) เฉลี่ย 8.01% โดยที่เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) สูงขึ้น 15.94% เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) สูงขึ้น 12.61% เทคนิควิธีเนออีฟเบย์ (Naive Bayes : NB) สูงขึ้น 8.23% เทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) สูงขึ้น 3.28% และเทคนิควิธีเคมีนส์ (K-means) มีค่าไม่เปลี่ยนแปลง แสดงดัง Figure 3 และ Figure 4

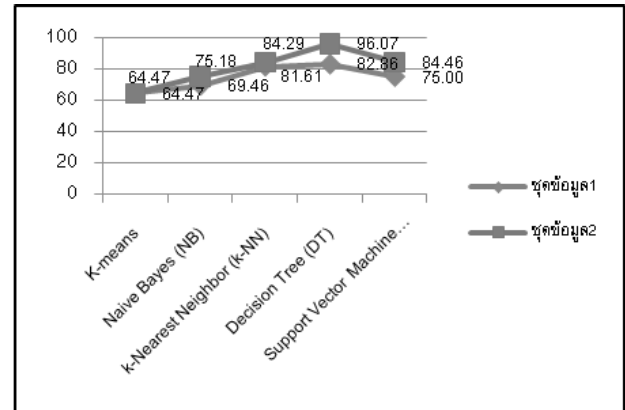


Figure 3 The chart shows the comparison of the accuracy values between data set 1 and data set 2

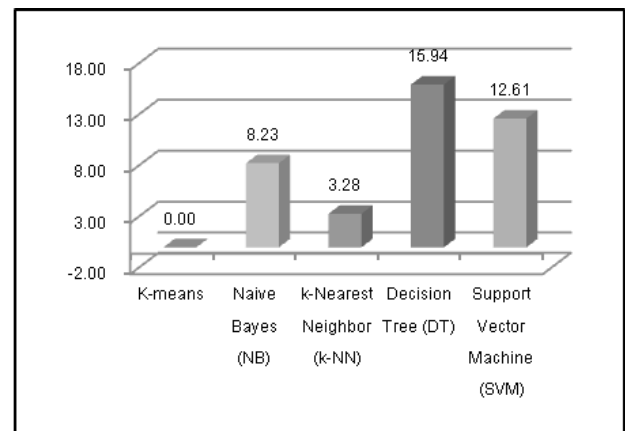


Figure 4 The chart shows the increment of the accuracy values between data set 1 and data set 2

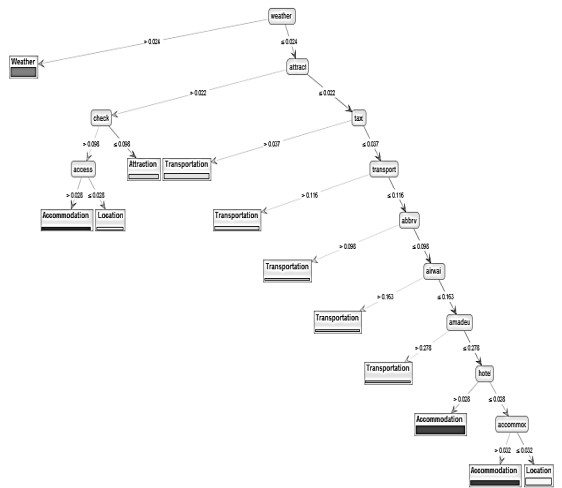


Figure 5 Web services classification model for Decision Tree of data set 1

Figure 5 แสดงโมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 1 (ข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย) ซึ่งให้ค่าความถูกต้อง (Accuracy) ในการจำแนกหมวดหมู่เว็บเซอร์วิสเท่ากับ 82.86%

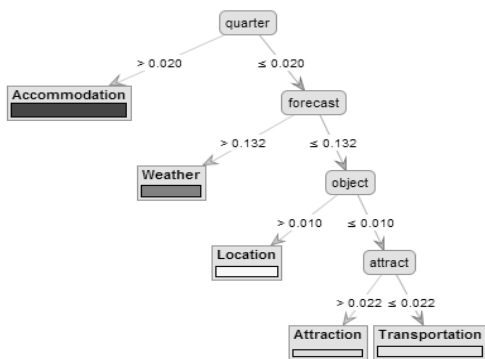


Figure 6 Web services classification model for Decision Tree of data set 2

Figure 6 แสดงโมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 2 (ข้อมูลของเว็บเซอร์วิสที่ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย) ซึ่งให้ค่าความถูกต้อง (Accuracy) ในการจำแนกหมวดหมู่เว็บเซอร์วิสเท่ากับ 96.07%

ผลการทดลองเมื่อนำมาเปรียบเทียบกับประสิทธิภาพกับผลงานวิจัยอื่น^{33,34,35} อาจไม่สามารถนำมาเปรียบเทียบกันได้โดยตรง เนื่องจากมีความแตกต่างของข้อมูลเว็บเซอร์วิสที่นำมาทดลองและกระบวนการในการคัดเลือกต่างกัน แต่ก็พบว่าผลการทดลองมีความสอดคล้องกันในเรื่องเทคนิควิธีการ

จำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้ค่าความถูกต้องสูงสุดคือเทคนิคต้นไม้ตัดสินใจและใช้แนวทางเชิงความหมายในการจำแนก

2. ผลการคัดเลือกเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยวและการประเมินประสิทธิภาพ

2.1 ส่วนติดต่อกับผู้ใช้ (User Interface)

ตัวอย่างส่วนติดต่อกับผู้ใช้ แสดงดัง Figure 7, Figure 8 และ Figure 9

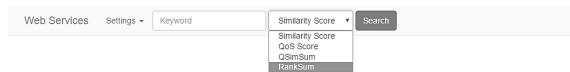


Figure 7 Home page of semantic web services searching

จาก Figure 7 ผู้ใช้สามารถกำหนดผลการเรียงลำดับของผลลัพธ์จากการค้นหา (User Preferences) ได้จากค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิส ค่าคะแนนความคล้ายคลึงกันเชิงความหมายระหว่างคำค้นจากผู้ใช้และข้อมูลในฐานความรู้แบบออนโทโลยี จากค่าเฉลี่ยของทั้งสองค่าข้างต้น และเรียงลำดับแบบวิธีหาผลรวม (RankSum)

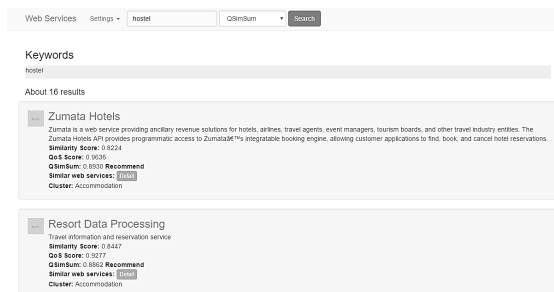


Figure 8 Sample results when searching with the term "hostel"

Web Service Name
 Description
 Operation
 Input
 Output

Submit

Web Service Similarity Table (Attraction)

	eContent_Maps	to.uri.st	Google Places	Rezgo
eContent.Maps	1.0000	0.7446	0.5000	0.8195
to.uri.st	0.7446	1.0000	0.5904	0.9117
Google Places	0.5000	0.5904	1.0000	0.6394
Rezgo	0.8195	0.9117	0.6394	1.0000

Figure 9 Shows the similarity values of Web Services in the Attraction group

2.2 ผลการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการ

1) ผลการประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการโดยการเปรียบเทียบผลลัพธ์ที่ได้รับระหว่างผู้เชี่ยวชาญและระบบสืบค้น

ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายโดยเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้น ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.84 ค่าความระลึก (Recall) เท่ากับ 0.80 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.82 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.93

2) ผลการประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการโดยใช้ไดนามิกเกณฑ์ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายโดยใช้ไดนามิกเกณฑ์ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.57 ค่าความระลึก (Recall) เท่ากับ 1.00 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.67 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.93

สรุปผลการวิจัยและข้อเสนอแนะ

1. สรุปผลการวิจัย

การจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวโดยใช้วิธีการทำเหมืองข้อมูล ได้มีการทดลองในหลายเทคนิควิธีเพื่อเปรียบเทียบว่าเทคนิคใดจะมีประสิทธิภาพและเหมาะสมกับข้อมูลเว็บเซอร์วิสที่ให้บริการที่เกี่ยวข้องด้านการท่องเที่ยวมากที่สุด ผลการประเมินผลและวัดประสิทธิภาพพบว่า เทคนิควิธีที่ให้ค่าความถูกต้อง (Accuracy) และประสิทธิภาพสูงที่สุดในการจำแนกหมวดหมู่ของเว็บเซอร์วิสคือเทคนิควิธีต้นไม้ตัดสินใจ ซึ่งสามารถนำโมเดลของเทคนิควิธีนี้ไปใช้ในการทำนายหรือจำแนกหมวดหมู่ของเว็บเซอร์วิสใหม่ที่ให้บริการที่เกี่ยวข้องด้านการท่องเที่ยวได้ นอกจากนี้ยังสามารถนำโมเดลนี้ไปเป็นแนวทางในการประยุกต์ใช้เพื่อจำแนกหมวดหมู่ของข้อมูลประเภทอื่น ๆ ได้

การพัฒนาฐานความรู้ออนโทโลยีเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวได้ประยุกต์จากออนโทโลยีที่เป็นมาตรฐานของดับเบิลยูทีซีที่มีชื่อว่า "Web Service Modeling Ontology: WSMO" จากนั้นได้ปรับปรุงให้สอดคล้องเพื่อรองรับกับข้อมูลด้านการท่องเที่ยว มีการจัดทำดัชนีของคำศัพท์ที่เกี่ยวข้องกับการท่องเที่ยวไว้ในคลังคำโดยใช้สคอส มีการสร้างความสัมพันธ์ของคำศัพท์เป็นโครงสร้างแบบลำดับชั้นตามรูปแบบอภิธานศัพท์มาตรฐานของเวิร์ดเน็ตซึ่งเป็นฐานข้อมูลคำศัพท์ภาษาอังกฤษขนาดใหญ่ และมีการเก็บรวบรวม

ข้อมูลเว็บเซอร์วิสจากหลายหลายเซอร์วิสที่จัดทำสร้างเป็นฐานความรู้แบบออนโทโลยีเพื่อให้สามารถอธิบายข้อมูลที่มีความเกี่ยวข้องในเชิงความหมายรองรับกับการสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายได้อย่างมีประสิทธิภาพ และเป็นต้นแบบของฐานความรู้ออนโทโลยีเว็บเซอร์วิสที่สามารถนำไปประยุกต์ใช้กับข้อมูลในโดเมนอื่น ๆ ได้

แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวหรือ SEMO ประกอบด้วย 1) การสร้างฐานความรู้โดยการออกแบบและพัฒนาออนโทโลยี ซึ่งประยุกต์จากออนโทโลยีเว็บเซอร์วิสที่ชื่อ Web Service Modeling Ontology (WSMO) ของดับเบิลยูทีซี และสร้างคลังคำโดยใช้สคอส 2) คุณภาพการบริการของเว็บเซอร์วิส มีการคำนวณหาค่าคะแนนคุณภาพของเว็บเซอร์วิส โดยพิจารณาจากปัจจัยคุณภาพที่สำคัญจำนวน 5 ปัจจัย ได้แก่ เวลาที่ตอบสนอง (Response Time) สภาพความพร้อมในการให้บริการ (Availability) ความเสถียรภาพหรือความน่าเชื่อถือ (Reliability) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) และความสำเร็จในการนำส่งข้อมูล (Throughput) สอดคล้องกับงานวิจัยของคาร์โดโซ และคณะ⁶ ซูซึลาและวาติเวล²⁰ เทาและเกวเจย์²³ ที่ได้ทำการศึกษาวิจัยเกี่ยวกับคุณภาพการให้บริการของเว็บเซอร์วิส พบว่ามีกรนำปัจจัยด้านคุณภาพมาใช้เป็นตัวกำหนดถึงคุณภาพการให้บริการ แต่จะแตกต่างกันไปตามบริบทและโดเมนของงานวิจัยนั้น และ 3) ระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมาย ผู้ใช้สามารถสืบค้นเว็บเซอร์วิสและสามารถกำหนดคุณลักษณะต่าง ๆ ในระบบสืบค้นให้เหมาะสมกับความต้องการของผู้ใช้ได้ เช่น กำหนดค่านำหนักของปัจจัยคุณภาพต่าง ๆ สอดคล้องกับงานวิจัยของเอจาและเดริส² แบดร์และคณะ⁴ โมฮิบิ ดิง และซี¹⁴ ที่ทำวิจัยเกี่ยวกับการค้นหาเว็บเซอร์วิสที่ผู้ใช้สามารถกำหนดคุณลักษณะในการสืบค้นที่ผู้ใช้ต้องการได้

2. ข้อเสนอแนะ

การปรับปรุงข้อมูลเว็บเซอร์วิสหรือคำศัพท์ในคลังคำในฐานความรู้แบบออนโทโลยีสามารถทำได้โดยอัตโนมัติ และสามารถปรับปรุงข้อมูลได้โดยผ่านทางโปรแกรมเว็บเบราว์เซอร์

การคำนวณหาค่าคุณภาพการบริการของเว็บเซอร์วิสในแต่ละกลุ่มเว็บเซอร์วิสสามารถทำได้โดยอัตโนมัติ โดยผู้ดูแลระบบสามารถปรับตั้งค่าให้อัพเดทข้อมูลได้อย่างอัตโนมัติตามวันเวลาหรือช่วงเวลาที่กำหนดเพื่อให้ค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิสที่ทันสมัยอยู่เสมอ ปรับปรุงส่วนติดต่อกับผู้ใช้ (User Interface) ที่ใช้ในการค้นหาเว็บเซอร์วิสเชิงความหมาย ให้สามารถรองรับ

การทำงานได้บนอุปกรณ์สมาร์ทโฟนและแท็บเล็ต รวมทั้งสามารถใช้งานได้ทั้งบนระบบปฏิบัติการแอนดรอยด์และไอโอเอส โดยทำเป็นโมบายแอปพลิเคชัน (Mobile Application) สามารถค้นหาเพื่อดาวน์โหลดและติดตั้งได้ง่ายผ่านทางเพย์สโตร์ (Play Store) หรือแอปสโตร์ (App Store) เพราะปัจจุบันอุปกรณ์ที่ผู้ใช้ทั่วโลกใช้ในการเข้าถึงอินเทอร์เน็ตเป็นอุปกรณ์เคลื่อนที่มากกว่าคอมพิวเตอร์ส่วนบุคคล (Cullen, A., www, 2016)

เอกสารอ้างอิง

1. Aghaei, S., Nematbakhsh M. A. and Farsani H. K. (2012). Evolution of the World Wide Web: From Web 1.0 to Web 4.0. *International Journal of Web & Semantic Technology (IJWesT)*, Vol. 3, No. 1, pp.1-10.
2. Ajao, T. A. and Deris, S. (2013). Optimal Web Service Selection with Consideration for User's Preferences. *International Journal of Computer Science Issues*, Vol.10, Issue 2, No.3, pp.355-359.
3. Al-Masri, E. and Mahmoud, Q. H. (2007). QoS-based Discovery and Ranking of Web Services. *IEEE 16th International Conference on Computer Communications and Networks (ICCCN '07)*, pp.529-534.
4. Badr, Y., Abraham, A., Biennier, F. and Grosan, C. (2008). Enhancing Web Service Selection by User Preferences of Non-functional Features. *Proceedings of the 2008 4th International Conference on Next Generation Web Services Practices*. pp.60-65.
5. Bruno, M., Canfora, G., Penta M.D. and Scognamiglio, R. (2005). An Approach to support Web Service Classification and Annotation. *Proceedings of the 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service*, pp.138-143.
6. Cardoso, J., Sheth, A., Miler, J., Arnold, J. and Kochut, K. (2004). Quality of Service for Workflows and Web Service Processes. *Journal of Web Semantics*, Vol. 1, pp.281-308.
7. Chaari, S., Badr, Y. and Biennier, F. (2008). Enhancing web service selection by QoS-based ontology and WS-policy. *Proceedings of the 2008 ACM symposium on Applied computing*, March 16-20, 2008, Fortaleza, Ceara, Brazil. pp.2426-2431.
8. Chang, G. (2012). QoS-Based Web Service Selection Approach. *Software Engineering and Knowledge Engineering: Theory and Practice. Advances in Intelligent and Soft Computing*, Vol. 2, Springer, Berlin, Heidelberg. pp.887-892.
9. Choudhury, N. (2014). World Wide Web and Its Journey from Web 1.0 to Web 4.0. *International Journal of Computer Science and Information Technologies*, Vol. 5 (6), pp.8096-8100.
10. Cullen, A. (2016). Internet Usage Worldwide: Mobile and tablet internet usage exceeds desktop for first time worldwide [On-line]. Available: <http://gs.statcounter.com/press/mobile-and-tablet-internet-usage-exceeds-desktop-for-first-time-worldwide>
11. Kuriakose, J. (2009). Understanding and Adopting Semantic Web Technology. *Cutter IT Journal. Cutter Information Corp.* 22 (9): pp.10-18.
12. Miao, D., Duan, Q., Zhang, H., and Jiao, N. (2009). Rough set based hybrid algorithm for text classification. *Expert Systems with Applications* 36(5): pp.9168-9174.
13. Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, Vol. 38, No. 11, pp.39-41.
14. Mohebi, A., Ding, C. and Chi, C-H. (2012). Efficient QoS-based Service Selection with Consideration of User Requirements. *2012 IEEE 16th International Enterprise Distributed Object Computing Conference*, pp.183-190.
15. Nayak, R. and Lee, B. (2007). Web service discovery with additional semantics and clustering. *Proceedings of the IEEE/WICI ACM International Conference on Web Intelligence*, pp.555-558.
16. Russell, J. and Roseby, P. (2006). *Dictionary of Leisure, Travel and Tourism* (3rd edition). A & C Black Publishers Ltd. United Kingdom.
17. Salton, G. (1988). *Automatic text processing*. Addison Wesley Longman Publishing Co., Inc., Boston, MA, USA.
18. Sidorov, G., Gelbukh, A., Gómez-Adorno, H., Pinto, D. (2014). Soft Similarity and Soft Cosine Measure: Similarity of Features in Vector Space Model. *Computación y Sistemas*. 18 (3): pp.491-504.

19. SKOS. (2009). Simple Knowledge Organization System Reference - W3C Recommendation [On-line]. Available: <http://www.w3.org/TR/skos-reference/>
20. Susila, and Vadivel, S. (2014). QoS Measurement Tool for Web Service Selection. *International Journal of Web Engineering* 2014, 3(1): pp.1-8.
21. Sycara, K., Paolucci, M., Soudry, J. and Srinivasan, N. (2004). Dynamic discovery and coordination of agent-based semantic web services. *IEEE Internet Computing*, Vol. 8, No. 3, pp.66-73.
22. Tabatabaei, S.G.H., Kadir, W.M.N. and Ibrahim, S. (2008). Semantic Web Service Discovery and Composition Based on AI Planning and Web Service Modeling Ontology. In *Proc. IEEE Asia-Pacific Services Computing Conference (APSCC '08)*, pp.387-403.
23. Tao, Y., and Kwei-Jay, L. (2004). Service selection algorithms for Web services with end-to-end QoS constraints. *e-Commerce Technology, 2004. CEC 2004, Proceedings IEEE International Conference*. pp.129-136.
24. UNWTO. (2015a). UNWTO Tourism Highlights (2015 Edition) [On-line]. Available: <http://www.e-unwto.org/doi/pdf/10.18111/9789284416899>
25. UNWTO. (2015b). Methodological Notes to the Tourism Statistics Database (2015 Edition). *Compendium of Tourism Statistics, Yearbook of Tourism Statistics, and Outbound tourism data*. p.21.
26. W3C-Semantic Web Activity. (2013). World Wide Web Consortium: Semantic Web Activity [On-line]. Available: <http://www.w3.org/2001/sw/>
27. W3C-SPARQL. (2013). World Wide Web Consortium: SPARQL 1.1 Overview - W3C Recommendation [On-line]. Available: <http://www.w3.org/TR/sparql11-overview/>
28. W3C-WS. (2004). World Wide Web Consortium: Web Services Architecture [On-line]. Available: <https://www.w3.org/TR/ws-arch/>
29. W3C-WSMO. (2005). World Wide Web Consortium: Web service modeling ontology (WSMO). W3C Member Submission [On-line]. Available: <http://www.w3.org/Submission/WSMO>
30. Wang, H., Shi, Y., Zhou, X., et al. (2010). Web Service Classification Using Support Vector Machine. *The 22nd IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2010)*, pp.3-6.
31. Wu, H. and Guo, C. (2011). The research and implementation of Web Service classification and discovery based on semantic. In *Computer Supported Cooperative Work in Design (CSCWD) 2011 15th IEEE International Conference*, pp.381-385.
32. Zhao, L., Ren, Y., Li, M. and Sakurai, K. (2012). Flexible service selection with user-specific QoS support in service-oriented architecture. *Journal of Network and Computer Applications*, Vol. 35, No. 3, pp.962-973.
33. Yang, J. and Zhou, X. (2015). Semi-automatic Web Service Classification Using Machine Learning. *International Journal of u- and e- Service, Science and Technology*, Vol. 8, No. 4, pp.339-348.
34. Sharma, S., Lather, J.S. and Dave, M. (2014). Semantic Approach for Classification of Web Services Using Unsupervised Normalized Similarity Measure. *Journal of Emerging Technologies in Web Intelligence*, Vol. 6, No. 3, pp.364-372.
35. Mohanty, R., Ravi, V. and Patra, M.R. (2010). Web-services classification using intelligent techniques. *International Journal: Expert Systems with Applications*, Vol. 37, No. 7, July 2010, pp.5484-5490.