แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว

นายเอกชัย แน่นอุดร

มหาวิทยาลัยเทคโนโลยีสุรนารี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาการสารสนเทศดุษฎีบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2559

A QOS-AWARE SEMANTIC WEB SERVICES SELECTION MODEL FOR TOURISM

Ekkachai Naenudorn

A Thesis Submitted in Partial Fulfillment of the Requirements for

the Degree of Doctor of Information Science in Information Technology

Suraneree University of Technology

Academic Year 2016

แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึง
คุณภาพการบริการสำหรับการท่องเที่ยว

มหาวิทยาลัยเทคโนโลยีสุรนารีอนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์

(รองศาสตราจารย์ ดร.วีรพงษ์ พลนิกรกิจ)
ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ศุภกฤษฎิ์ นิวัฒนากูล)
กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)

(ผู้ช่วยศาสตราจารย์ ดร.ภุชงค์ อุทโยภาศ)
กรรมการ

(รองศาสตราจารย์ ดร.นิตยา เกิดประสพ)
กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ธรา อั่งสกุล)
กรรมการ

(ศาสตราจารย์ ดร.ชูกิจ ลิมปิจำนงค์)
รองอธิการบดีฝ่ายวิชาการและนวัตกรรม

(รองศาสตราจารย์ ดร.วีรพงษ์ พลนิกรกิจ)
คณบดีสำนักวิชาเทคโนโลยีสังคม

มหาวิทยาลัยเทคโนโลยีสุรนารี

เอกชัย แน่นอุดร : แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว (A QOS-AWARE SEMANTIC WEB SERVICES SELECTION MODEL FOR TOURISM) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.ศุภกฤษฎิ์ นิวัฒนากูล, 140 หน้า.

งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว และเพื่อวิเคราะห์เปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว ปัจจัยที่ใช้ในการพิจารณาถึงคุณภาพการบริการของเว็บเซอร์วิส ได้แก่ เวลาตอบสนอง สภาพความพร้อม ความเสถียรภาพ ค่าบริการ และความสำเร็จในการนำส่งข้อมูล กระบวนการในการดำเนินการวิจัยประกอบไปด้วย การสร้างฐานความรู้ออนโทโลยีของเว็บเซอร์วิสในโดเมนด้านการท่องเที่ยว การจำแนกหมวดหมู่ของเว็บเซอร์วิสโดยใช้หลักการทำเหมืองข้อมูลและเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิส การวัดความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่ม การคำนวณหาค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิส การสืบค้นเว็บเซอร์วิสโดยใช้วิธีการถ่วงค่าน้ำหนักของคำ และการค้นหาเว็บเซอร์วิสที่ผู้ใช้สามารถกำหนดเงื่อนไขและค่าน้ำหนักให้กับปัจจัยคุณภาพต่าง ๆ ได้ วิธีการทั้งหมดนี้จะทำให้การคัดเลือกเว็บเซอร์วิสเชิงความหมายมีประสิทธิภาพ และผลลัพธ์ที่ได้จากการค้นหาเว็บเซอร์วิสผู้ใช้จะได้รับเว็บเซอร์วิสที่มีคุณภาพและตรงกับความต้องการของผู้ใช้มากที่สุด

ผลการประเมินผลและวัดประสิทธิภาพโดยใช้วิธีการทำเหมืองข้อมูลในการจำแนกหมวดหมู่ของเว็บเซอร์วิส พบว่าเทคนิควิธีที่ให้ค่าความถูกต้อง (Accuracy) สูงที่สุดคือเทคนิควิธีต้นไม้ตัดสินใจ มีค่าเท่ากับ 96.07% รองลงมาคือเทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน มีค่าเท่ากับ 84.46% เทคนิควิธีเค์เนียร์เรสท์เน่เบอร์ มีค่าเท่ากับ 84.29% เทคนิควิธีเนอีฟเบย์ มีค่าเท่ากับ 75.18% และเทคนิควิธีเคมีนส์ มีค่าเท่ากับ 64.47% ตามลำดับ

ผลการประเมินผลและวัดประสิทธิภาพของระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยว พบว่าผลการประสิทธิภาพโดยการเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้น ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.8391 ค่าความระลึก (Recall) เท่ากับ 0.8018 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.8175 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.9270 และผลการประเมินประสิทธิภาพโดยใช้ไดนามิกเกณฑ์ ให้ค่าความแม่นยำเท่ากับ 0.5660 ค่าความระลึกเท่ากับ 1.0000 ค่าเอฟเมเชอร์เท่ากับ 0.6720 และค่าความถูกต้องเท่ากับ 0.9306

สาขาวิชาเทคโนโลยีสารสนเทศ ลายมือชื่อนักศึกษา _____

ปีการศึกษา 2559 ลายมือชื่ออาจารย์ที่ปรึกษา _____

EKKACHAI NAENUDORN : A QOS-AWARE SEMANTIC WEB SERVICES SELECTION MODEL FOR TOURISM. THESIS ADVISOR : ASST.PROF. SUPHAKIT NIWATTANAKUL, Ph.D., 140 PP.

WEB SERVICES SELECTION / SEMANTIC WEB SERVICE / ONTOLOGY / DATA MINING / CLASSIFICATION / QUALITY OF SERVICE

This research aims to design and develop a QoS-aware semantic web services selection model for tourism. And, 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, classification web services using data mining principles and techniques, measuring the similarity of web services in each group, calculating web services quality scores, using the term weighting method, searching web services where users can define terms and weight values for various quality factors. All of these mentioned methods will make the selection of semantic web services effiectively and the result of web service search will enable users to get the best quality of the web service that truely meets their needs.

Evaluation results and performance measurements using data mining methods for classification web services found that the highest accuracy technique was 96.07% Decision Tree technique, 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 service according to the quality of service for tourism were that, by comparing between experts and search systems, Precision was 0.8391, Recall was 0. 8018, F-measure was 0. 8175, and Accuracy was 0. 9270. The performance evaluation using dynamic criteria for Precision was 0.5660, Recall was 1.0000, F-measure was 0.6720 and Accuracy was 0.9306.

มหาวิทยาลัยเทคโนโลยีสุรนารี

School of Information Technology Student's Signature Ekkachai

Academic Year 2016 Advisor's Signature S. Ninet

กิตติกรรมประกาศ

ความสำเร็จในครั้งนี้จะเกิดขึ้นไม่ได้ ถ้าไม่ได้รับคำปรึกษา แนะนำ ช่วยเหลือ สนับสนุน และกำลังใจ จากผู้มีพระคุณดังต่อไปนี้

ขอขอบพระคุณอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร.ศุภกฤษฎิ์ นิวัฒนากูล และอาจารย์ผู้ได้ให้คำปรึกษาแนะนำ อาจารย์ ดร.นิศาชล จำนงศรี ที่คอยให้คำปรึกษา แนะนำ และช่วยผลักดันให้งานวิจัยนี้สำเร็จไปด้วยดี

ขอขอบคุณ รองศาสตราจารย์ ดร.วีรพงษ์ พลนิกรกิจ ประธานกรรมการสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.ภุชงค์ อุทโยภาศ รองศาสตราจารย์ ดร.นิตยา เกิดประสพ และผู้ช่วยศาสตราจารย์ ดร.ธรา อั่งสกุล คณะกรรมการสอบวิทยานิพนธ์ ที่กรุณาให้คำแนะนำปรึกษา และตรวจทานเล่มวิทยานิพนธ์จนเสร็จสมบูรณ์

ขอขอบคุณ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม ที่ให้การสนับสนุนทุนการศึกษา จนสำเร็จการศึกษาตามหลักสูตรนี้

ขอขอบคุณ บุคลากรประจำสาขาวิชาเทคโนโลยีสารสนเทศ และบุคลากรประจำสำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี ที่ให้ความช่วยเหลือและอำนวยความสะดวกต่าง ๆ ตลอดระยะเวลาที่ศึกษานี้

ขอขอบคุณ เพื่อนร่วมหลักสูตรวิทยาการสารสนเทศดุษฎีบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศทุกท่าน โดยเฉพาะเพื่อนร่วมรุ่น (รุ่นที่ 2 รหัส 54) และขอบคุณน้องโอห์ม วิทวัส ที่คอยให้ความช่วยเหลือและให้กำลังใจจนสำเร็จการศึกษา

สุดท้ายนี้ขอกราบขอบพระคุณบิดา มารดา ขอบคุณที่เป็นกำลังใจและเป็นแรงบันดาลใจที่สำคัญจนทำให้ผู้วิจัยประสบความสำเร็จในการศึกษาในครั้งนี้

เอกชัย แน่นอุดร

สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย) ..... ก
บทคัดย่อ (ภาษาอังกฤษ) ..... ข
กิตติกรรมประกาศ ..... ง
สารบัญ ..... จ
สารบัญตาราง ..... ซ
สารบัญรูป ..... ฎ

บทที่

1 บทนำ 1
1.1 ความสำคัญและที่มาของปัญหาการวิจัย 1
1.2 วัตถุประสงค์ 5
1.3 คำถามนำการวิจัย 5
1.4 ข้อตกลงเบื้องต้น 5
1.5 ขอบเขตของการวิจัย 5
1.6 ประโยชน์ที่คาดว่าจะได้รับ 5
1.7 คำอธิบายศัพท์ 6
2 ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง 7
2.1 ความหมายและความสำคัญของการท่องเที่ยว 7
2.2 เว็บเซอร์วิสด้านการท่องเที่ยว 9
2.2.1 ความหมายของเว็บเซอร์วิส 9
2.2.2 กระบวนการทำงานของเว็บเซอร์วิส 9
2.2.3 การบริการเว็บเซอร์วิส 10
2.2.4 เทคโนโลยีที่เกี่ยวข้องกับเว็บเซอร์วิส 10
2.3 การค้นหาข้อมูลเชิงความหมาย 18
2.4 ออนโทโลยีและเทคโนโลยีที่เกี่ยวข้อง 20

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญ (ต่อ)

หน้า

2.4.1 ทฤษฎีพื้นฐานเกี่ยวกับออนโทโลยี 20
2.4.2 องค์ประกอบของออนโทโลยี 21
2.4.3 กระบวนการในการพัฒนาออนโทโลยี 22
2.4.4 งานวิจัยที่เกี่ยวข้องกับออนโทโลยี 22
2.4.5 สคอส (SKOS) 24
2.4.6 เวิร์ดเน็ต (WordNet) 24
2.4.7 ภาษาสปาร์เคิล (SPARQL) 25
2.5 การคัดเลือกเว็บเซอร์วิสเชิงความหมาย (Semantic Web Services Selection) 26
2.6 คุณภาพการให้บริการของเว็บเซอร์วิส 28
2.6.1 ปัจจัยที่ใช้กำหนดคุณภาพการให้บริการของเว็บเซอร์วิส 28
2.6.2 ปัจจัยคุณภาพการให้บริการของเว็บเซอร์วิส 29
2.6.3 การคำนวณค่าคุณภาพการให้บริการของเว็บเซอร์วิส 32
2.7 การทำเหมืองข้อมูล (Data Mining) 33
2.7.1 ความหมายของการทำเหมืองข้อมูล 33
2.7.2 ขั้นตอนการทำเหมืองข้อมูล 34
2.7.3 ลักษณะงานของการทำเหมืองข้อมูล 35
2.7.4 การจำแนกหมวดหมู่ของเว็บเซอร์วิสโดยใช้วิธีการทำเหมืองข้อมูล 36
2.8 เทคนิคการวัดความคล้ายคลึงกันแบบโคไซน์ (Cosine Similarity) 40
2.9 การจัดลำดับข้อมูล (Ranking) 42
2.10 งานวิจัยที่เกี่ยวข้อง 43
3 วิธีดำเนินการวิจัย 47
3.1 วิธีวิจัย 48
3.1.1 ศึกษาทฤษฎีและเทคนิคการคัดเลือกเว็บเซอร์วิส 48
3.1.2 การออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมาย โดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว 49
3.2 เครื่องมือที่ใช้ในการวิจัย 86
3.3 การวิเคราะห์ข้อมูล 86

สารบัญ (ต่อ)

หน้า

4 ผลการวิจัยและอภิปรายผล 88
4.1 ผลการจำแนกหมวดหมู่ของเว็บเซอร์วิสและการประเมินประสิทธิภาพ 88
4.2 ผลการคัดเลือกเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยวและการประเมินประสิทธิภาพ 97
5 สรุปผลการวิจัยและข้อเสนอแนะ 110
5.1 สรุปผลการวิจัย 110
5.2 อภิปรายผลการวิจัย 111
5.3 ข้อเสนอแนะ 113
รายการอ้างอิง 114
ภาคผนวก 121
ประวัติผู้เขียน 140

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญตาราง

ตารางที่ **หน้า**

2.1 ส่วนประกอบของเอกสารดับเบิ้ลยูเอสดีแอล 14
2.2 สรุปงานวิจัยเกี่ยวกับปัจจัยด้านคุณภาพการให้บริการ (Quality of Service: QoS) 29
2.3 สรุปงานวิจัยที่เกี่ยวข้องด้านการคัดเลือกเว็บเซอร์วิส (Web Service Selection) 45
3.1 รายละเอียดของออนโทโลยีเว็บเซอร์วิส 51
3.2 ตัวอย่างคำศัพท์บางส่วนในคลังคำในฐานความรู้ออนโทโลยี 55
3.3 ตัวอย่างผลลัพธ์ที่ได้จากการทำความสะอาดข้อมูล (Data Cleaning) ให้กับชุดสตริง 58
3.4 ตัวอย่างผลลัพธ์ที่ได้จากการตัดคำหยุด (Stop words) 59
3.5 ตัวอย่างผลลัพธ์ที่ได้จากการหารากศัพท์ (Stemming Word) 60
3.6 ตัวอย่างผลลัพธ์ที่ได้จากการเปรียบเทียบระหว่างคำศัพท์กับคลังคำในออนโทโลยี 61
3.7 การแทนค่าเพื่อคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส รอบที่ 1 65
3.8 การแทนค่าเพื่อคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส รอบที่ 2 65
3.9 ตัวอย่างการคำนวณหาค่าเวลาที่ตอบสนอง 69
3.10 ตัวอย่างการคำนวณหาค่าสภาพความพร้อมในการให้บริการ 69
3.11 ตัวอย่างการคำนวณหาค่าความเสถียรของระบบ 70
3.12 ตัวอย่างการคำนวณหาค่า ค่าใช้จ่ายในการเรียกใช้บริการ 71
3.13 ตัวอย่างการคำนวณหาค่าความสำเร็จในการนำส่งข้อมูล 72
3.14 ค่าคะแนนคุณภาพการให้บริการของเว็บเซอร์วิสหลังถ่วงน้ำหนักแล้ว 73
3.15 ตัวอย่างการให้ค่าน้ำหนักของคำตามโครงสร้างสคอส (SKOS) 74
3.16 ตัวอย่างการกำหนดค่าน้ำหนักของคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส 75
3.17 ตัวอย่างค่าคะแนน Similarity และค่าคะแนน QoS 79
3.18 ค่าน้ำหนัก Similarity และค่าน้ำหนัก QoS 79
3.19 ผลลัพธ์ของค่าคะแนน QsimSum ของแต่ละเว็บเซอร์วิส 79
3.20 ตัวอย่างค่าคะแนน Similarity และค่าคะแนน QoS 80
3.21 ค่าลำดับของแต่ละเว็บเซอร์วิส 80
3.22 ค่าส่วนกลับลำดับของ Similarity และ QoS 80

สารบัญตาราง (ต่อ)

ตารางที่ **หน้า**

3.23 ค่าคะแนนของ Similarity และ QoS 80
3.24 ค่าคะแนน Ranksum ก่อนถ่วงน้ำหนัก 80
3.25 ค่าคะแนน Ranksum 81
3.26 คอนฟิวส์ชันเมทริกซ์ (Confusion Matrix) ของผลการสืบค้น 82
4.1 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคมีนส์ (K-means) 89
4.2 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) 89
4.3 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) 89
4.4 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) 90
4.5 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) 90
4.6 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคมีนส์ (K-means) 90
4.7 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) 91
4.8 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) 91
4.9 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) 91
4.10 เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) 91
4.11 สรุปผลการวัดประสิทธิภาพการจัดกลุ่มเว็บเซอร์วิส แสดงค่าความแม่นยำ ค่าความระลึก ค่าเอฟเมเชอร์ และค่าความถูกต้อง ของชุดข้อมูลในแบบที่ 1 (ข้อมูลของเว็บเซอร์วิสที่-ไม่ได้ผ่านการหาความสัมพันธ์ของคำ) 92
4.12 สรุปผลการวัดประสิทธิภาพการจัดกลุ่มเว็บเซอร์วิส แสดงค่าความแม่นยำ ค่าความระลึก ค่าเอฟเมเชอร์ และค่าความถูกต้อง ของชุดข้อมูลในแบบที่ 2 (ข้อมูลของเว็บเซอร์วิสที่-ผ่านการหาความสัมพันธ์ของคำเชิงความหมายในออนโทโลยี) 93
4.13 ค่าเอฟเมเชอร์ (F-measure) และค่าเบี่ยงเบนมาตรฐาน (Standard Devision : SD) 98
4.14 ตัวอย่างคำสำคัญที่ใช้ในการประเมินผลการสืบค้นเว็บเซอร์วิส 103

สารบัญตาราง (ต่อ)

ตารางที่ **หน้า**

4.15 จำนวนเว็บเซอร์วิสที่ถูกค้นพบเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้น แยกตามคำค้น/คำสำคัญ และแสดงค่า Truth Table 104
4.16 ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น 104
4.17 ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น 105
4.18 ผลการประเมินผลการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น 106
4.19 จำนวนเว็บเซอร์วิสที่ถูกค้นพบโดยใช้ไดนามิกเกณฑ์ และแสดงค่า Truth Table 107
4.20 ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์ 107
4.21 ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์ 108
4.22 ผลการประเมินผลการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์ 109

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญรูป

รูปที่ **หน้า**

2.1 แนวโน้มและการคาดการณ์เกี่ยวกับการท่องเที่ยวโลก ตั้งแต่ปี ค.ศ. 1950 ถึงปี ค.ศ. 2030 8
2.2 แบบจำลองของเว็บเซอร์วิส 10
2.3 ตัวอย่างภาษาเอ็กซ์เอ็มแอล 11
2.4 เว็บเซอร์วิสในรูปแบบโซพ (Simple Object Access Protocol: SOAP) 12
2.5 โครงสร้างของเอกสารโซพ (SOAP) 13
2.6 ตัวอย่างเอกสาร SOAP request 13
2.7 ตัวอย่างเอกสาร SOAP response 13
2.8 ตัวอย่างโครงสร้างของเอกสารดับเบิ้ลยูเอสดีแอล (WSDL) 15
2.9 รูปแบบโครงสร้างการทำงานของเรสท์ (REpresentational State Transfer: REST) 17
2.10 โครงสร้างความสัมพันธ์ของคำศัพท์แบบสคอส 24
2.11 ตัวอย่างภาษาสปาร์เคิล (SPARQL) ในการค้นหาข้อมูลในออนโทโลยี 26
2.12 ค่าความคล้ายคลึงแบบโคไซน์ที่แสดงถึงความแตกต่างระหว่างสองเอกสาร 41
3.1 กรอบแนวคิดการวิจัย 47
3.2 สถาปัตยกรรมของระบบ SEMO 49
3.3 บางส่วนของออนโทโลยีเว็บเซอร์วิส 51
3.4 แสดงออนโทโลยีเว็บเซอร์วิสในโปรแกรมโพรเทเจ 53
3.5 พจนานุกรมด้านการสันทนาการ การเดินทางและการท่องเที่ยว 53
3.6 เว็บไซต์อภิธานศัพท์เวิร์ดเน็ตออนไลน์สำหรับค้นหาคำศัพท์ที่เกี่ยวข้อง 54
3.7 ตัวอย่างคลังคำโดยใช้สคอสในโปรแกรมโพรเทเจ 54
3.8 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “hostel” 56
3.9 ผลลัพธ์จากการสอบถามข้อมูล (Query) ชื่อของเว็บเซอร์วิสและคำอธิบาย 57
3.10 ผลลัพธ์จากการสอบถามข้อมูล Interface, Operation และ Output ของเว็บเซอร์วิส 58
3.11 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “hotel” 60
3.12 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “reservation” 61
3.13 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Accommodation 66

สารบัญรูป (ต่อ)

รูปที่ **หน้า**

3.14 อัลกอริทึมการคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส 67
3.15 โปรแกรมอะโดบีดรีมเวฟเวอร์สำหรับออกแบบส่วนติดต่อกับผู้ใช้ 74
3.16 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “hostel” (SKOS Vocabulary Relationship) 75
3.17 สถาปัตยกรรมของระบบค้นหาเว็บเซอร์วิสเชิงความหมาย 81
4.1 แผนภูมิแสดงค่าความแม่นยำ ค่าความระลึก และค่าเอฟเมเชอร์ของชุดข้อมูลในแบบที่ 1 92
4.2 แผนภูมิแสดงค่าความถูกต้องของชุดข้อมูลในแบบที่ 1 93
4.3 แผนภูมิแสดงค่าความแม่นยำ ค่าความระลึก และค่าเอฟเมเชอร์ของชุดข้อมูลในแบบที่ 2 94
4.4 แผนภูมิแสดงค่าความถูกต้องของชุดข้อมูลในแบบที่ 2 94
4.5 แผนภูมิแสดงการเปรียบเทียบค่าความถูกต้องระหว่างข้อมูลชุดที่ 1 และข้อมูลชุดที่ 2 95
4.6 แผนภูมิแสดงอัตราผลต่างของค่าความถูกต้องระหว่างข้อมูลชุดที่ 1 และข้อมูลชุดที่ 2 96
4.7 โมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 1 96
4.8 โมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 2 97
4.9 แผนภูมิแสดงค่าเอฟเมเชอร์ (F-measure) 98
4.10 หน้าแรกของระบบการค้นหาเว็บเซอร์วิสเชิงความหมาย 99
4.11 ตัวอย่างผลลัพธ์เมื่อค้นหาด้วยคำค้น “hostel” 99
4.12 หน้าการตั้งค่าน้ำหนัก (Weight Configuration) 100
4.13 หน้าการคำนวณหาค่าคะแนนคุณภาพการบริการของแต่ละกลุ่มเว็บเซอร์วิส 100
4.14 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Accommodation 101
4.15 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Location 101
4.16 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Transportation 102
4.17 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Attraction 102
4.18 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Weather 103
4.19 ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมาย ระหว่างผู้เชี่ยวชาญและระบบสืบค้น 106
4.20 ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมาย โดยใช้ไดนามิกเกณฑ์ 109

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของปัญหาการวิจัย

องค์การการท่องเที่ยวโลกแห่งสหประชาชาติหรือ UNWTO (United Nations World Tourism Organization) ได้คาดการณ์ว่าอนาคตการท่องเที่ยวโลกไปจนถึงปี ค.ศ. 2030 จะเติบโตอย่างต่อเนื่อง โดยตั้งแต่ปี ค.ศ. 1995 เป็นต้นมา จำนวนนักท่องเที่ยวทั่วโลกที่คาดการณ์ไว้ค่อนข้างเป็นไปตามเป้าหมาย แต่ในบางปีอาจมีจำนวนต่ำกว่าที่คาดการณ์ไว้เนื่องจากวิกฤตเศรษฐกิจและภัยสงคราม และในอนาคตจำนวนนักท่องเที่ยวระหว่างประเทศจะเพิ่มขึ้นโดยเฉลี่ยปีละ 43 ล้านคน โดยตั้งแต่ปี ค.ศ. 2015 จำนวนนักท่องเที่ยวจากเขตเศรษฐกิจเกิดใหม่จะมีจำนวนมากกว่านักท่องเที่ยวจากประเทศพัฒนาแล้ว โดยภูมิภาคเอเชียแปซิฟิกจะได้รับนักท่องเที่ยวจากตลาดเกิดใหม่มากที่สุด (UNWTO, 2015a) เมื่อการท่องเที่ยวโลกมีแนวโน้มเติบโตมากขึ้นในอนาคต ทำให้มีความต้องการข้อมูลเกี่ยวกับการท่องเที่ยวมากขึ้นตามไปด้วย เช่น ข้อมูลเกี่ยวกับสถานที่ท่องเที่ยว ตำแหน่งที่ตั้ง แผนที่ ข้อมูลเกี่ยวกับระบบคมนาคม และที่พัก เป็นต้น ดังนั้นการนำเสนอข้อมูลด้านการท่องเที่ยวที่มีความถูกต้อง ทันสมัย และรวดเร็วจึงมีความสำคัญ ส่งผลทำให้บุคคลทั่วไปสนใจการท่องเที่ยวมากขึ้นและทำให้เศรษฐกิจด้านการท่องเที่ยวของโลกเติบโตขึ้น (สำนักงานปลัดกระทรวงการท่องเที่ยวและกีฬา, 2559)

การนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์สามารถแบ่งตามยุคของเว็บได้เป็น 4 ยุค (Aghaei et al., 2012; Choudhury, 2014) ดังนี้

เว็บในยุคที่ 1 เป็นยุคที่เรียกว่าสแตติกเว็บ (Static Web) ซึ่งเป็นจุดเริ่มต้นของเทคโนโลยีเว็บเกิดขึ้นเมื่อปี ค.ศ.1994 โดยผู้เริ่มต้นคือ ทิม เบอร์เนอร์ลี (Tim Berners-Lee) ได้ก่อตั้งองค์กรเว็บไซต์สากล (World Wide Web Consortium: W3C) ขึ้นที่สถาบันเทคโนโลยีแมสซาชูเซตส์ เพื่อสร้างมาตรฐานและข้อเสนอแนะสำหรับใช้เป็นหลักในการพัฒนาคุณภาพของเว็บไซต์ การนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ในยุคนี้ใช้ภาษาหลักในการพัฒนาเว็บไซต์คือภาษาเอชทีเอ็มแอล (Hyper Text Markup Language: HTML) เป็นการออกแบบเว็บไซต์ที่มีการนำเสนอข้อมูลแบบตายตัว เหมาะกับการนำเสนอข้อมูลที่ไม่มีการเปลี่ยนแปลงบ่อยนัก การจัดการเนื้อหาทำได้ไม่ค่อยสะดวก

เว็บในยุคที่ 2 เป็นยุคที่เรียกว่าไดนามิกเว็บ (Dynamic Web) ซึ่งแต่ละองค์กรจะพัฒนาเป็นเว็บแอปพลิเคชันเพื่อนำเสนอข้อมูลด้านการท่องเที่ยว ผู้ใช้สามารถโต้ตอบกับเว็บไซต์ได้ เช่น

การโพสต์ข้อความ ในยุคนี้ข้อมูลจะถูกจัดเก็บไว้ในฐานข้อมูล ทำให้การจัดการข้อมูลทำได้สะดวก มีระบบจัดการข้อมูลหลังร้าน (Back Office) แต่ผู้ดูแลระบบจะต้องเตรียมข้อมูลเองและปรับปรุงข้อมูลให้มีความทันสมัยอยู่ตลอดเวลา การนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ในยุคนี้ใช้ภาษาสคริปต์ในการพัฒนา เช่น PHP, ASP, ASP.NET, JSP, Javascript เป็นต้น ส่วนฐานข้อมูลที่นิยมใช้ เช่น Oracle, Microsoft SQL Server, MySQL, MongoDB, PostgreSQL เป็นต้น

เว็บในยุคที่ 3 เป็นยุคที่เรียกว่าเว็บเซอร์วิส (Web Service) (W3C-WS, www, 2004) เป็นการพัฒนาเป็นเว็บแอปพลิเคชันเพื่อเรียกใช้บริการเว็บเซอร์วิสที่ให้บริการข้อมูลด้านการท่องเที่ยว เว็บเซอร์วิสเป็นเทคโนโลยีที่ทำให้โปรแกรมประยุกต์ต่าง ๆ ที่พัฒนาขึ้นโดยมีพื้นฐานทางด้านภาษาและสถาปัตยกรรมที่แตกต่างกัน สามารถสื่อสารแลกเปลี่ยนข้อมูลกันได้โดยใช้ภาษาเอ็กซ์เอ็มแอล (Extensible Markup Language: XML) (W3C-XML, www, 2000) เป็นภาษากลางในการสื่อสารแลกเปลี่ยนข้อมูลโดยอาศัยโพรโทคอลเอชทีทีพี (Hypertext Transfer Protocol: HTTP) และใช้ร่วมกับมาตรฐานเกี่ยวกับเว็บอื่น ๆ เว็บเซอร์วิสใช้ภาษาดับเบิ้ลยูเอสดีแอล (Web Services Description Language: WSDL) (W3C-WSDL, www, 2007) ในการอธิบายถึงรูปแบบการบริการข้อมูล การใช้งานเว็บเซอร์วิสปัจจุบันอาจอยู่ในรูปแบบของโซพ (Simple Object Access Protocol: SOAP) (W3C-SOAP, www, 2007) ตามแบบมาตรฐานแบบดั้งเดิมหรือตามแนวทางแบบเรสท์ (Representational State Transfer: REST) (Fielding, 2000) ที่เรียบง่ายไม่ซับซ้อน การนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ หลายองค์กรเลือกที่จะใช้ข้อมูลด้านการท่องเที่ยวที่ได้รับจากการให้บริการของผู้ให้บริการเว็บเซอร์วิส ซึ่งมีความเป็นมืออาชีพและมีความเชี่ยวชาญเฉพาะด้าน ทำให้ได้รับข้อมูลที่ถูกต้องและทันสมัย ซึ่งบางบริการฟรีแต่บางบริการมีการเก็บค่าธรรมเนียมในการใช้บริการ

เว็บในยุคที่ 4 ของการนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ เป็นยุคที่เรียกว่าซีแมนติกเว็บ (Semantic Web) หรือเว็บเชิงความหมาย (Kuriakose, 2009; W3C-Semantic, www, 2013) เป็นพัฒนาการของเวิลด์ไวด์เว็บที่เกี่ยวข้องกับข้อมูลและการบริการบนเว็บไซต์ ในการสร้างความเป็นไปได้ที่เว็บไซต์จะสามารถเข้าใจถึงความต้องการของผู้ใช้ เว็บเชิงความหมายเป็นวิธีการที่จะทำให้คอมพิวเตอร์หรือโปรแกรมประยุกต์สามารถเข้าใจข้อมูลที่สอดคล้องกับความเข้าใจของมนุษย์ สามารถเข้าใจความหมายของคำและแนวความคิดที่รวบรวมเกี่ยวกับคำในขอบเขตของความรู้ที่ต้องการตามที่ผู้พัฒนาได้กำหนดไว้โดยเชื่อมโยงความสัมพันธ์ของข้อมูลแบบมีโครงสร้าง เว็บเชิงความหมายจึงเป็นอีกแนวคิดหนึ่งที่ถูกนำไปประยุกต์ใช้ในการพัฒนาระบบการค้นคืนสารสนเทศ (Information Retrieval: IR) ในรูปแบบที่เรียกว่า การสืบค้นเชิงความหมาย (Semantic Search) ซึ่งเป็นวิธีการค้นหาทรัพยากรสารสนเทศที่ต้องการจากคลังสารสนเทศภายใต้เงื่อนไขหรือกระบวนการ

คัดเลือกที่ผู้ใช้ต้องการ ทำให้สามารถค้นหาข้อมูลได้ถูกต้องครอบคลุมถึงข้อมูลที่มีความสัมพันธ์เชิงความหมายและได้ข้อมูลที่ตรงตามความต้องการของผู้ใช้มากขึ้น

จากข้อมูลการนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยวบนเว็บไซต์ทั้ง 4 ยุค สรุปได้ว่าเทคโนโลยีเว็บเซอร์วิสและการสืบค้นข้อมูลเชิงความหมายมีความความสำคัญต่อการนำไปใช้ในการนำเสนอข้อมูลเกี่ยวกับการท่องเที่ยว เทคโนโลยีเว็บเซอร์วิสทำให้ผู้ใช้สามารถออกแบบและพัฒนาโปรแกรมประยุกต์ที่ไม่จำกัดแพลตฟอร์มเพื่อเรียกใช้เว็บเซอร์วิสจากผู้ให้บริการที่เป็นมืออาชีพและมีความเชี่ยวชาญเฉพาะด้าน ทำให้ได้รับข้อมูลที่ถูกต้องและทันสมัย ส่วนการสืบค้นข้อมูลเชิงความหมายจะทำให้ผู้ใช้สามารถค้นหาเว็บเซอร์วิสได้ตรงตามความต้องการของผู้ใช้มากขึ้น

การค้นหาเว็บเซอร์วิสในปัจจุบันผู้ใช้สามารถค้นหาได้จากยูดีดีไอรีจิสตรี (UDDI registry) ฟังก์ชันการทำงานของแต่ละเว็บเซอร์วิสถูกประกาศไว้ในไฟล์ดับเบิ้ลยูเอสดีแอล (WSDL) ซึ่งผู้ให้บริการได้ประกาศไว้บนยูดีดีไอรีจิสตรี ลักษณะของการค้นหาเว็บเซอร์วิสเป็นการใช้คำสำคัญในการค้นหา โดยการนำคำสำคัญไปเปรียบเทียบกับข้อมูลเอกสารเว็บเซอร์วิสที่เก็บในฐานข้อมูล (Keyword Matching) ผลลัพธ์ที่ได้จะพบเฉพาะรายการของเว็บเซอร์วิสที่มีคำสำคัญตรงกันซึ่งอาจไม่ถูกต้องและครบถ้วน เพราะบางเว็บเซอร์วิสอาจมีความสัมพันธ์กันถึงแม้ว่าเว็บเซอร์วิสนั้นอาจจะไม่มีคำสำคัญนั้นปรากฏในเอกสารเว็บเซอร์วิสเลยก็ตาม ความสัมพันธ์ในลักษณะนี้เรียกว่า ความสัมพันธ์เชิงความหมาย ดังนั้นการสืบค้นเชิงความหมายจึงเป็นแนวทางหนึ่งที่จะช่วยให้การค้นหาข้อมูลมีประสิทธิภาพมากขึ้นเพราะสามารถเข้าใจความหมายที่แท้จริงของเอกสาร

เว็บเซอร์วิสที่ให้บริการข้อมูลด้านการท่องเที่ยวได้รับความสนใจจากผู้ใช้มากขึ้น ดังนั้นจึงมีงานวิจัยที่เกี่ยวข้องกับการคัดเลือกเว็บเซอร์วิสเกิดขึ้นหลายงานวิจัยแต่ยังมีหลายประเด็นที่ควรปรับปรุงเพื่อให้การคัดเลือกเว็บเซอร์วิสเชิงความหมายมีประสิทธิภาพมากขึ้น โดยสามารถสรุปประเด็นปัญหาที่สำคัญของงานวิจัยที่เกี่ยวข้องกับการคัดเลือกเว็บเซอร์วิสเชิงความหมายได้ดังนี้

1) งานวิจัยที่เกี่ยวข้องกับการคัดเลือกเว็บเซอร์วิสส่วนใหญ่จะพิจารณาถึงปัจจัยที่มีผลต่อคุณภาพการให้บริการของเว็บเซอร์วิส แต่ไม่ได้ให้ความสำคัญในเรื่องการนำเสนอเว็บเซอร์วิสที่ตรงกับความต้องการของผู้ใช้ (Badr et al., 2008; Mohebi et al., 2012; Ajao and Deris, 2013; Zhao et al., 2012)

2) การคัดเลือกเว็บเซอร์วิสจะพิจารณาคัดเลือกเว็บเซอร์วิสทั้งหมดที่มีอยู่ในฐานข้อมูล ทำให้กระบวนการในการคัดเลือกใช้เวลานาน ซึ่งความเป็นจริงแล้วเว็บเซอร์วิสนั้นมีอยู่เป็นจำนวนมากและมีหลายกลุ่มบริการ ดังนั้นก่อนทำการพิจารณาคัดเลือกด้วยเกณฑ์หรือเทคนิควิธีต่าง ๆ ควรจำแนกหมวดหมู่ของเว็บเซอร์วิสเพื่อให้การคัดเลือกเว็บเซอร์วิสมีประสิทธิภาพมากขึ้น (Bruno et al., 2005; Wang et al., 2010; Wu and Guo, 2011)

3) ลักษณะของการค้นหาเว็บเซอร์วิสในปัจจุบันจะใช้คำสำคัญจากผู้ใช้ไปเปรียบเทียบกับข้อมูลเอกสารเว็บเซอร์วิส (Keyword Matching) ผลลัพธ์ที่ได้อาจไม่ถูกต้องและครบถ้วน (Sycara et al., 2004; Nayak and Lee, 2007; Tabatabaei et al., 2008)

4) เว็บเซอร์วิสที่มีให้บริการบนเครือข่ายอินเทอร์เน็ตในปัจจุบันมีจำนวนมากและหลากหลายกลุ่มบริการ จะทราบได้อย่างไรว่าเว็บเซอร์วิสใดในแต่ละกลุ่มบริการมีคุณภาพการบริการดีที่สุด หรือในกรณีที่บางเว็บเซอร์วิสใช้งานไม่ได้ จะมีเว็บเซอร์วิสใดที่สามารถใช้ทดแทนกันได้ (Menasce, 2002; Liu et al., 2004; Al-Masri and Mahmoud, 2007; Chaari et al., 2008; D'Mello et al., 2010; Chang, 2012; Wang et al., 2013)

ดังนั้นการวิจัยในครั้งนี้จึงได้ทำการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการ โดยเน้นไปที่ข้อมูลเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว งานวิจัยนี้ประยุกต์จากข้อมูลกลุ่มการท่องเที่ยวในวารสารขององค์การการท่องเที่ยวโลก (UNWTO, 2015b) โดยสามารถแบ่งได้เป็น 5 กลุ่ม ได้แก่ ที่พัก (Accommodation) การเดินทางขนส่ง (Transportation) สถานที่ท่องเที่ยว (Attraction) ตำแหน่งที่ตั้ง (Location) และสภาพอากาศ (Weather) งานวิจัยนี้ใช้การทำเหมืองข้อมูล (Data Mining) โดยใช้เทคนิคการจำแนกข้อมูล (Classification) มาใช้ในการจำแนกหมวดหมู่ของเว็บเซอร์วิส มีการพิจารณาถึงคุณภาพการบริการของเว็บเซอร์วิส โดยพิจารณาจาก 5 ปัจจัย ได้แก่ เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ (Response Time) สภาพความพร้อมในการให้บริการ (Availability) ความเสถียรภาพหรือความน่าเชื่อถือ (Reliability) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) และความสำเร็จในการนำส่งข้อมูล (Throughput) งานวิจัยนี้สร้างฐานความรู้โดยใช้ออนโทโลยีของเว็บเซอร์วิสในโดเมนด้านการท่องเที่ยวในการอธิบายความหมายของข้อมูล มีการจำแนกหมวดหมู่ของเว็บเซอร์วิสก่อนที่จะทำการวัดความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่ม จากนั้นวิเคราะห์และเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสเพื่อหาเทคนิควิธีที่มีประสิทธิภาพมากที่สุด มีการคำนวณหาค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิส ขั้นตอนการสืบค้นใช้วิธีการถ่วงค่าน้ำหนักของคำ และในการค้นหาเว็บเซอร์วิสผู้ใช้สามารถกำหนดเงื่อนไขและกำหนดค่าน้ำหนักให้กับปัจจัยคุณภาพต่าง ๆ ได้อย่างอิสระ (User Preferences) วิธีการทั้งหมดนี้จะทำให้การคัดเลือกเว็บเซอร์วิสเชิงความหมายมีประสิทธิภาพ และผลลัพธ์ที่ได้จากการค้นหาเว็บเซอร์วิสผู้ใช้จะได้รับเว็บเซอร์วิสที่มีคุณภาพและตรงกับความต้องการของผู้ใช้มากที่สุด

1.2 วัตถุประสงค์

1.2.1 เพื่อออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว

1.2.2 เพื่อวิเคราะห์และเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่เกี่ยวข้องกับการบริการด้านการท่องเที่ยว

1.3 คำถามนำการวิจัย

1.3.1 แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวควรมีกระบวนการและวิธีการคัดเลือกเว็บเซอร์วิสอย่างไร

1.3.2 แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวควรใช้เทคนิคใดในการจำแนกหมวดหมู่ของเว็บเซอร์วิส

1.4 ข้อตกลงเบื้องต้น

ผู้ใช้ต้องระบุคำค้นเป็นภาษาอังกฤษที่สะกดถูกต้องตามพจนานุกรมเข้าสู่แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว

1.5 ขอบเขตของการวิจัย

งานวิจัยนี้เป็นการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการ โดยมุ่งเน้นไปที่ข้อมูลเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว ประกอบไปด้วย 5 กลุ่ม ได้แก่ ที่พัก (Accommodation) การเดินทางขนส่ง (Transportation) สถานที่ท่องเที่ยว (Attraction) ตำแหน่งที่ตั้ง (Location) และสภาพอากาศ (Weather) ซึ่งเป็นกลุ่มที่ได้จากการประยุกต์จากกลุ่มการท่องเที่ยวในวารสารขององค์การการท่องเที่ยวโลกแห่งสหประชาชาติหรือ UNWTO (United Nations World Tourism Organization) (UNWTO, 2015b)

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1.6.1 ได้แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว

1.6.2 ได้เทคนิคที่เหมาะสมในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการที่เกี่ยวข้องกับการท่องเที่ยว

1.7 คำอธิบายศัพท์

1.7.1 การคัดเลือกเว็บเซอร์วิสเชิงความหมาย หมายถึง กระบวนการที่จะทำการคัดเลือกเว็บเซอร์วิสจากฐานความรู้แบบออนโทโลยีที่มีความเกี่ยวข้องสัมพันธ์กันกับคำสำคัญที่รับมาจากผู้ใช้ ซึ่งการคัดเลือกเว็บเซอร์วิสเชิงความหมายประกอบไปด้วยหลายขั้นตอน เช่น สร้างฐานความรู้โดยใช้ออนโทโลยีของเว็บเซอร์วิสในโดเมนด้านการท่องเที่ยว การจำแนกหมวดหมู่ของเว็บเซอร์วิส การคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส การถ่วงน้ำหนักของคำ การคำนวณหาค่าคุณภาพของเว็บเซอร์วิส และการค้นหาเว็บเซอร์วิสเชิงความหมาย เป็นต้น

1.7.2 การค้นหาเว็บเซอร์วิสเชิงความหมาย หมายถึง การสืบค้นหรือค้นคืนข้อมูลเว็บเซอร์วิสในฐานความรู้แบบออนโทโลยีที่มีการจัดเก็บและเชื่อมโยงข้อมูลเว็บเซอร์วิสต่าง ๆ ที่มีความเกี่ยวข้องเข้าไว้ด้วยกันอย่างมีโครงสร้างและเป็นลำดับชั้น โดยระบบจะสามารถเข้าใจความหมายที่แท้จริงที่มีความสัมพันธ์กันระหว่างข้อมูลในฐานความรู้แบบออนโทโลยีและคำสำคัญจากผู้ใช้ ความสัมพันธ์ในลักษณะนี้เรียกว่า ความสัมพันธ์เชิงความหมาย ซึ่งจะช่วยให้การค้นหาข้อมูลมีประสิทธิภาพมากขึ้น

1.7.3 คุณภาพการบริการของเว็บเซอร์วิส หมายถึง เกณฑ์ที่ใช้กำหนดถึงประสิทธิภาพการให้บริการของเว็บเซอร์วิส ปัจจัยด้านคุณภาพการบริการต่าง ๆ เช่น เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ (Response Time) สภาพความพร้อมในการให้บริการ (Availability) ความเสถียรภาพหรือความน่าเชื่อถือของระบบ (Reliability) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) และความสำเร็จในการนำส่งข้อมูล (Throughput) เป็นต้น

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 2
ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

งานวิจัยเรื่องแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวในครั้งนี้ได้ศึกษาถึงวรรณกรรมและงานวิจัยที่เกี่ยวข้อง เพื่อประมวลแนวคิดทฤษฎีและเทคนิควิธีต่าง ๆ เพื่อใช้ในการทำวิจัย มีรายละเอียดดังต่อไปนี้

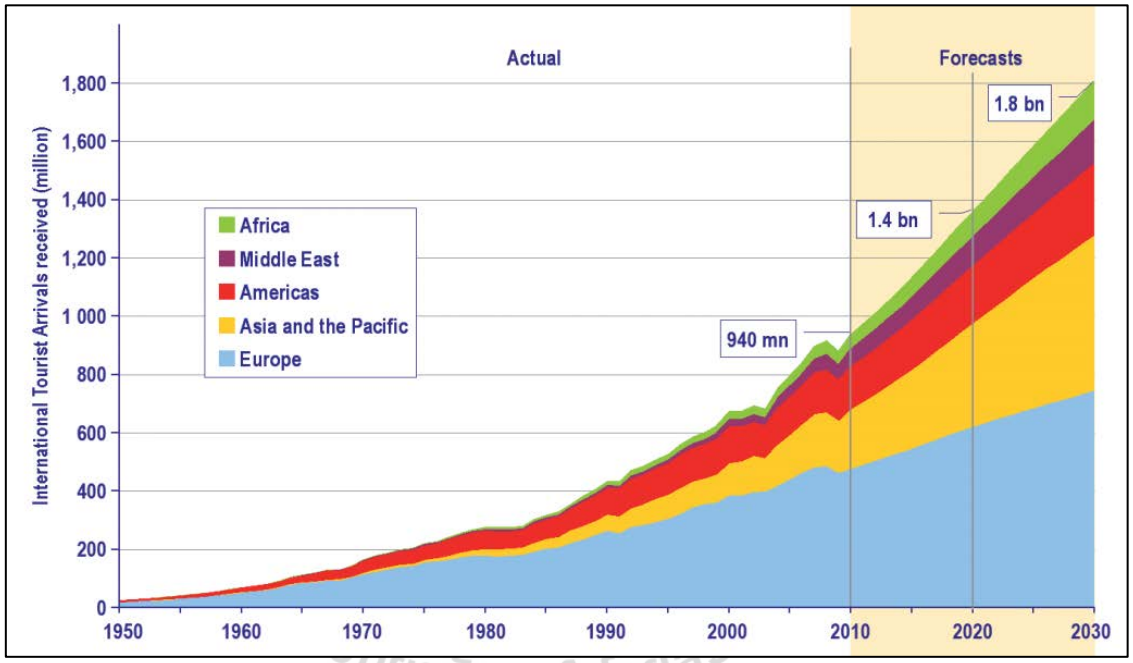
2.1 ความหมายและความสำคัญของการท่องเที่ยว

องค์การการท่องเที่ยวโลก (World Tourism Organization หรือ UNWTO) ให้ความหมายของการท่องเที่ยว ว่าหมายถึง การเดินทางใด ๆ ที่เป็นการเดินทางตามเงื่อนไขสากล 3 ประการ คือ 1) การเดินทางจากที่อยู่อาศัยปกติไปยังที่อื่นเป็นการชั่วคราว แต่ไม่ใช่ไปตั้งหลักแหล่งเป็นการถาวร 2) การเดินทางนั้นเป็นไปด้วยความสมัครใจ หรือความพึงพอใจของผู้เดินทางเอง ไม่ใช่เป็นการถูกบังคับ ไม่ใช่เพื่อทำสงคราม และ 3) เป็นการเดินทางด้วยวัตถุประสงค์ใด ๆ ก็ตามที่ไม่ใช่ประกอบอาชีพหรือหารายได้

UNWTO ได้คาดการณ์อนาคตการท่องเที่ยวโลกไปจนถึงปี ค.ศ. 2030 ไว้ว่า การเติบโตของจำนวนนักท่องเที่ยวทั่วโลกที่ UNWTO คาดการณ์ไว้จนถึงปี ค.ศ. 2020 ค่อนข้างเป็นไปตามเป้าหมาย คือมีการเติบโตอย่างต่อเนื่องตั้งแต่ปี ค.ศ. 1995-2020 แต่ในบางปีอาจมีจำนวนต่ำกว่าที่คาดการณ์ไว้เนื่องจากวิกฤตเศรษฐกิจและภัยสงคราม และในอนาคต UNWTO ได้คาดการณ์ไว้ว่าจำนวนนักท่องเที่ยวระหว่างประเทศจะเพิ่มขึ้นโดยเฉลี่ยปีละ 43 ล้านคน โดยคาดการณ์จำนวนนักท่องเที่ยวจะเพิ่มเป็น 1.4 พันล้านคนในปี ค.ศ. 2020 และ 1.8 พันล้านคนในปี ค.ศ. 2030 โดยตั้งแต่ปี ค.ศ. 2015 เป็นต้นไป จำนวนนักท่องเที่ยวจากเขตเศรษฐกิจเกิดใหม่จะมีจำนวนมากกว่านักท่องเที่ยวจากประเทศพัฒนาแล้ว โดยภูมิภาคเอเชียแปซิฟิกจะได้รับนักท่องเที่ยวจากตลาดเกิดใหม่มากที่สุด (UNWTO, 2015b)

โอกาสและความท้าทายที่จะเกิดขึ้นในอนาคต UNWTO คาดการณ์ไว้ 5 ประการ ดังนี้ 1) การท่องเที่ยวจะเป็นเครื่องมือที่ทรงอำนาจในการพัฒนาเศรษฐกิจและสังคม รวมทั้งลดความยากจนในหลายประเทศจากการสร้างงาน การพัฒนาโครงสร้างพื้นฐานและรายได้จากการส่งออก 2) เพื่อก้าวไปสู่ศักยภาพนี้จำเป็นต้องสร้างและยกระดับการรับรู้ และทำให้การท่องเที่ยวเป็นเรื่องสำคัญในวาระทางการเมือง 3) ความยั่งยืนทางสังคม เศรษฐกิจ และสิ่งแวดล้อม มีความสำคัญมากขึ้น โดยเน้นเรื่อง

การพึ่งพาพลังงาน การปรับตัวจากการเปลี่ยนแปลงของสภาพภูมิอากาศ เศรษฐกิจสีเขียว การจัดการพื้นที่แออัด และการบริหารความเสี่ยง 4) ความต้องการของผู้บริโภคที่เปลี่ยนแปลงง่ายเนื่องจากมีประสบการณ์มากขึ้น มีความต้องการสูงขึ้น รวมทั้งการเปลี่ยนแปลงทางประชากรศาสตร์ การเปลี่ยนแปลงค่านิยม และวิถีชีวิตจากเศรษฐกิจเชิงบริการสู่เศรษฐกิจเชิงประสบการณ์ และ 5) การแข่งขันที่เพิ่มขึ้นในแง่ของการสร้างสภาพแวดล้อมทางธุรกิจที่เหมาะสม เช่น นวัตกรรม การพัฒนาผลิตภัณฑ์ ความหลากหลายของสินค้า ตลาดและกลุ่มลูกค้า การตลาดและการส่งเสริมสินค้า การใช้เทคโนโลยีสารสนเทศ และการพัฒนาทรัพยากรบุคคล เป็นต้น



รูปที่ 2.1 แนวโน้มและการคาดการณ์เกี่ยวกับการท่องเที่ยวโลก ตั้งแต่ปี ค.ศ. 1950 ถึงปี ค.ศ. 2030 (UNWTO, 2015b)

เมื่อการท่องเที่ยวโลกมีแนวโน้มเติบโตมากขึ้นในอนาคต ทำให้มีความต้องการข้อมูลเกี่ยวกับการท่องเที่ยวมากขึ้นตามไปด้วย ดังนั้นงานวิจัยนี้จึงมุ่งเน้นไปที่ข้อมูลของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวซึ่งเป็นข้อมูลที่สำคัญและสามารถใช้ประโยชน์ได้ในอนาคต ข้อมูลของกลุ่มที่เกี่ยวข้องกับการท่องเที่ยวในงานวิจัยนี้ประยุกต์จากกลุ่มการท่องเที่ยวในวารสารขององค์การการท่องเที่ยวโลก (UNWTO, 2015a) โดยสามารถแบ่งได้ 5 กลุ่ม ได้แก่ ที่พัก (Accommodation) การเดินทางขนส่ง (Transportation) สถานที่ท่องเที่ยว (Attraction) ตำแหน่งที่ตั้ง (Location) และสภาพอากาศ (Weather)

2.2 เว็บเซอร์วิสด้านการท่องเที่ยว

2.2.1 ความหมายของเว็บเซอร์วิส

เว็บเซอร์วิสเป็นเทคโนโลยีที่ทำให้โปรแกรมประยุกต์ต่าง ๆ ที่พัฒนาขึ้นโดยมีพื้นฐานทางด้านภาษาและสถาปัตยกรรมที่แตกต่างกัน สามารถสื่อสารแลกเปลี่ยนข้อมูลกันได้โดยใช้ภาษาเอ็กซ์เอ็มแอล (W3C-XML, www, 2002) เป็นภาษากลางในการสื่อสารแลกเปลี่ยนข้อมูลโดยอาศัยโพรโทคอลเอชทีทีพี (Hypertext Transfer Protocol: HTTP) และใช้ร่วมกับมาตรฐานเกี่ยวกับเว็บอื่น ๆ เว็บเซอร์วิสใช้ภาษาดับเบิ้ลยูเอสดีแอล (Web Services Description Language: WSDL) (W3C-WSDL, www, 2007) ในการอธิบายถึงรูปแบบการบริการข้อมูล การใช้งานเว็บเซอร์วิสปัจจุบันอาจอยู่ในรูปแบบของโซพ (Simple Object Access Protocol: SOAP) (W3C-SOAP, www, 2007) ตามแบบมาตรฐานแบบดั้งเดิมหรือตามแนวทางแบบเรสท์ (Representational State Transfer: REST) (Fielding, 2000) ที่เรียบง่ายไม่ซับซ้อน โดยมีหน่วยงานหลักที่ดูแลมาตรฐานและสถาปัตยกรรมของเว็บเซอร์วิสคือโอเอซิส (OASIS) และดับเบิ้ลยูทรีซี (W3C) เว็บเซอร์วิสสามารถทำงานร่วมกันได้กับทุกแฟลตฟอร์ม (Platform Independent) หรือสามารถทำงานได้บนระบบปฏิบัติการที่แตกต่างกันได้ และเว็บเซอร์วิสสามารถทำงานร่วมกันได้กับทุกภาษาโปรแกรม

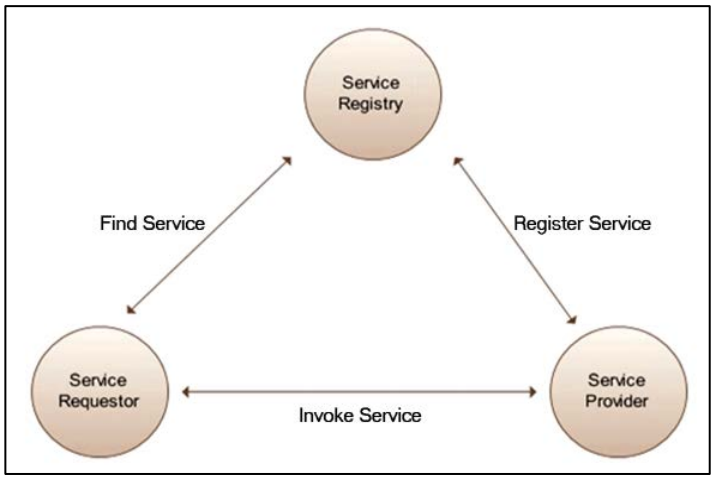
2.2.2 กระบวนการทำงานของเว็บเซอร์วิส

กระบวนการการทำงานของเว็บเซอร์วิสจะมีขั้นตอนการทำงานตามแนวคิดแบบสถาปัตยกรรมเชิงบริการ (Service Oriented Architecture: SOA) ซึ่งสามารถแบ่งองค์ประกอบของเว็บเซอร์วิสได้เป็น 3 ส่วน ดังแสดงในรูปที่ 2.2 แบบจำลองของเว็บเซอร์วิส และสามารถอธิบายได้ดังนี้

1) ผู้ให้บริการ (Service Provider) มีหน้าที่ในการพัฒนาและติดตั้งเว็บเซอร์วิส เป็นผู้ที่นิยามความหมายของเซอร์วิสและลงทะเบียนเว็บเซอร์วิสกับเซอร์วิสรีจิสทรี (Service Registry) เพื่ออธิบายถึงการบริการเว็บเซอร์วิสของตนและประกาศบริการสู่สาธารณะ

2) ผู้ใช้บริการ (Service Requestor) เป็นผู้เรียกใช้เว็บเซอร์วิส โดยทำการค้นหาเซอร์วิสจากเซอร์วิสไดเรกทอรี แล้วทำการเรียกใช้เซอร์วิสจากผู้ให้บริการ

3) เซอร์วิสรีจิสทรี (Service Registry) มีหน้าที่ในการรับลงทะเบียนและช่วยในการค้นหาเว็บเซอร์วิส Service Registry จะเก็บรายละเอียดของเว็บเซอร์วิสต่าง ๆ เช่น นิยาม และตำแหน่งของเว็บเซอร์วิส ทำหน้าที่คล้ายกับสมุดโทรศัพท์เพื่อช่วยให้ผู้ใช้บริการสามารถค้นหาเซอร์วิสที่ต้องการได้ ซึ่งฟังก์ชันการทำงานของเว็บเซอร์วิสถูกประกาศไว้ใน WSDL ซึ่งสามารถค้นหาได้จากเซอร์วิสรีจิสทรีนี้



รูปที่ 2.2 แบบจำลองของเว็บเซอร์วิส

2.2.3 การบริการเว็บเซอร์วิส

การบริการของเว็บเซอร์วิสมีหลายรูปแบบ แต่รูปแบบที่เป็นที่นิยมกันคือโซพ (Simple Object Access Protocol: SOAP) (W3C-SOAP, www, 2007) และเรสท์ (Representational State Transfer: REST) (Fielding, 2000) เว็บเซอร์วิสแบบโซพ (SOAP) จะเป็นรูปแบบมาตรฐานดั้งเดิมตั้งแต่ยุคเริ่มต้นของเว็บเซอร์วิส ส่วนเว็บเซอร์วิสแบบเรสท์ (REST) เป็นมาตรฐานใหม่ที่กำลังเป็นที่นิยม ซึ่งรายละเอียดของเว็บเซอร์วิสทั้งสองรูปแบบอธิบายไว้ในหัวข้อถัดไป

เว็บเซอร์วิสที่ให้บริการในประเทศไทย เช่น เว็บเซอร์วิสที่ให้บริการเกี่ยวกับข้อมูลราคาน้ำมันของ ปตท. เว็บเซอร์วิสของกรมสรรพากร เว็บเซอร์วิสของศูนย์ปฏิบัติการจังหวัดเชียงใหม่ เว็บเซอร์วิสที่ให้บริการเกี่ยวกับข้อมูลการท่องเที่ยวของจังหวัดขอนแก่น เป็นต้น ส่วนเว็บเซอร์วิสของต่างประเทศ เช่น Amazon Web services, Google Maps API, Yahoo APIs and Web Services และ xmethods เป็นต้น

2.2.4 เทคโนโลยีที่เกี่ยวข้องกับเว็บเซอร์วิส

1) เอ็กซ์เอ็มแอล (Extensible Markup Language: XML) (W3C-XML, www, 2002)

เอ็กซ์เอ็มแอลเป็นภาษามาตรฐานที่ถูกออกแบบมาเพื่อให้เป็นภาษากลางสำหรับแลกเปลี่ยนข้อมูล มีลักษณะข้อมูลเป็น text-based เป็นภาษาที่ไม่ติดอยู่กับแพลตฟอร์มใดๆ ดังนั้นระบบที่อยู่ต่างแพลตฟอร์มกัน สามารถแลกเปลี่ยนข้อมูลกันได้ โดยผู้ใช้เองสามารถกำหนดแท็ก (Tag) ที่ใช้สื่อความหมายของเนื้อหา กำหนดโครงสร้างของเอกสาร และยังสามารถนำเอ็กซ์เอ็มแอลนั้นไปประมวลผลต่อในแอปพลิเคชันอื่นได้ ในปัจจุบันภาษาเอ็กซ์เอ็มแอลได้กลายเป็นภาษามาตรฐานสำคัญสำหรับการกำหนดโครงสร้างข้อมูล เนื้อหา และรูปแบบของข้อมูลของเอกสาร

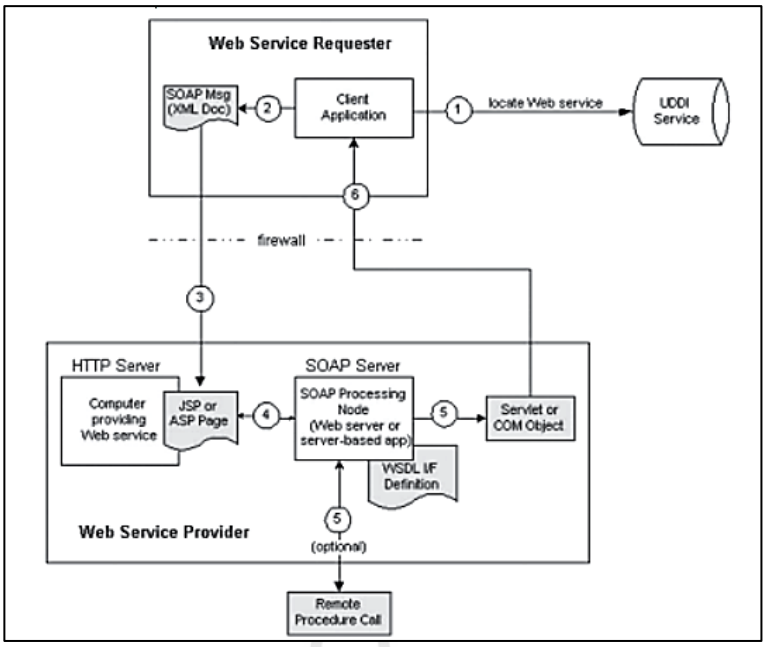
อิเล็กทรอนิกส์ และยังมีการพัฒนาเพื่อให้สามารถแลกเปลี่ยนข้อมูลระหว่างหน่วยงาน โปรแกรม ประยุกต์ และอุปกรณ์ต่าง ๆ ผ่านทางอินเทอร์เน็ตได้อีกด้วย

```
<?xml version="1.0" ?>
<?xml:stylesheet?>
<purchase-order>
  <header>
    <po-number>1234</po-number>
    <date>1999-02-08</date><time>14:05</time>
  </header>
  <billing>
    <company>XMLSolutions</company>
    <address>
      <street>601 Pennsylvania Ave. NW</street>
      <street>Suite 900</street>
      <city>Washington</city>
      <st>DC</st><postcode>20004</postcode>
    </address>
  </billing>
  <order items="1">
    <item>
      <reference>097251</reference>
      <description>Widgets</description>
      <quantity>4</quantity>
      <unit-price>11.99</unit-price>
      <price>47.46</price>
    </item>
    <tax type="sales">
      <tax-unit>VA</tax-unit>
      <calculation>0.045</calculation>
      <amount>2.16</amount>
    </tax>
    ...
  </order>
</purchase-order>
```

รูปที่ 2.3 ตัวอย่างภาษาเอ็กซ์เอ็มแอล

2) โซพ (Simple Object Access Protocol: SOAP) (W3C-SOAP, www, 2007)

โซพคือโพรโทคอลในการสื่อสารข้อมูลระหว่างเว็บเซอร์วิส มีรากฐานมาจากภาษาเอ็กซ์เอ็มแอล (XML) ทำให้เว็บเซอร์วิสสามารถสื่อสารกันได้ถึงแม้ว่าคอมพิวเตอร์จะมีความแตกต่างกันทางด้านแพลตฟอร์ม และทำงานอยู่บนโพรโทคอลเอชทีทีพี การส่งข้อความของโซพมีสองรูปแบบคือ SOAP-RPC และ SOAP message โดย SOAP-RPC ใช้ในการส่งข้อความเพื่อใช้เรียกเมธอดหรือ Procedure ซึ่งโดยมากจะเป็นรูปแบบสมมาตร (Synchronous) โดยโซพจะส่ง SOAP Request และข้อมูลต่างๆ เพื่อเรียกใช้เมธอดในการประมวลผล และจะรอให้ได้ผลลัพธ์การประมวลผลที่ส่งกลับมาแบบ SOAP Response ส่วน SOAP-message ใช้ในการส่งข่าวสารหรือข้อมูลในรูปแบบเอ็กซ์เอ็มแอล (XML) ระหว่างผู้ให้บริการและผู้ใช้บริการ โดยสามารถส่งได้ทั้งแบบสมมาตร (Synchronous) และแบบไม่สมมาตร (Asynchronous)



รูปที่ 2.4 เว็บเซอร์วิสในรูปแบบโซพ (Simple Object Access Protocol: SOAP)

จากรูปที่ 2.4 เป็นการทำงานของเว็บเซอร์วิสที่ใช้โซพ (SOAP) สามารถอธิบายเป็นขั้นตอนต่าง ๆ ได้ดังนี้

1. ค้นหาบริการที่ต้องการจากไดเรกทอรีกลางหรือยูดีดีไอ (Universal Description, Discovery and Integration: UDDI) อ่านข้อกำหนดของการใช้บริการจาก WSDL โดยดาวน์โหลดจากยูดีดีไอหรือจะดาวน์โหลดจากผู้ให้บริการเว็บเซอร์วิสโดยตรง

2. สร้างคำร้องในรูปแบบของโซพ (SOAP) ตามข้อกำหนดในดับเบิ้ลยูเอสดีแอล (WSDL)

3. ส่งคำร้องไปยังเครื่องแม่ข่ายผู้ให้บริการ

4. เครื่องแม่ข่ายประมวลผล

5. เครื่องแม่ข่ายมีการเรียกใช้วัตถุภายใน

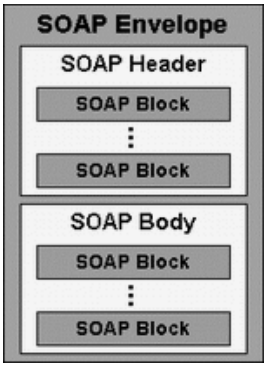
6. เครื่องแม่ข่ายสร้างคำตอบในรูปแบบของโซพแล้วส่งกลับ

โครงสร้างของเอกสารโซพ (SOAP) แบ่งได้เป็น 3 ส่วนหลักดังนี้คือ

1. SOAP Envelop สำหรับใช้ในการอธิบายและระบุเนื้อหา (Content) ของเอกสารทั้งหมด

2. SOAP Header ส่วนเพิ่มเติมของเอกสารโซพ ซึ่งจะมีก็ได้หรือไม่มีก็ได้

3. SOAP Body ส่วนที่ใช้ในการเรียกใช้งานเซอร์วิสและผลลัพธ์ที่ได้



รูปที่ 2.5 โครงสร้างของเอกสารโซพ (SOAP)

```
POST /InStock HTTP/1.1
Host: www.example.org
Content-Type: application/soap+xml; charset=utf-8
Content-Length: nnn

<?xml version="1.0"?>
<soap:Envelope
xmlns:soap="http://www.w3.org/2001/12/soap-envelope"
soap:encodingStyle="http://www.w3.org/2001/12/soap-encoding">

<soap:Body xmlns:m="http://www.example.org/stock">
  <m:GetStockPrice>
    <m:StockName>IBM</m:StockName>
  </m:GetStockPrice>
</soap:Body>

</soap:Envelope>
```

รูปที่ 2.6 ตัวอย่างเอกสาร SOAP request

```
HTTP/1.1 200 OK
Content-Type: application/soap+xml; charset=utf-8
Content-Length: nnn

<?xml version="1.0"?>
<soap:Envelope
xmlns:soap="http://www.w3.org/2001/12/soap-envelope"
soap:encodingStyle="http://www.w3.org/2001/12/soap-encoding">

<soap:Body xmlns:m="http://www.example.org/stock">
  <m:GetStockPriceResponse>
    <m:Price>89.5</m:Price>
  </m:GetStockPriceResponse>
</soap:Body>

</soap:Envelope>
```

รูปที่ 2.7 ตัวอย่างเอกสาร SOAP response

3) ดับเบิ้ลยูเอสดีแอล (Web Services Description Language: WSDL) (W3C-WSDL, www, 2007)

ดับเบิ้ลยูเอสดีแอลคือภาษาที่อธิบายรายละเอียดในการติดต่อกับเว็บเซอร์วิส เพื่อให้ทราบถึงวิธีการเรียกใช้งานเว็บเซอร์วิส โครงสร้างของภาษามีรากฐานมาจากภาษาเอ็กซ์เอ็มแอล โดยจะแบ่งการอธิบายเว็บเซอร์วิสออกเป็นสองส่วน คือส่วนที่เป็นนามธรรม (Abstract) เพื่ออธิบายโอเปอเรชัน (Operation) อินพุตและเอาท์พุตพารามิเตอร์ และส่วนที่เป็นรูปธรรม (Concrete) เพื่ออธิบายโพรโทคอลของเน็ตเวิร์ค ตำแหน่งของจุดปลายทาง (Endpoint Address) และรูปแบบของข้อมูล ปัจจุบันดับเบิ้ลยูเอสดีแอลเป็นภาษาที่อยู่ในความดูแลของดับเบิ้ลยูทรีซี (W3C) สำหรับเวอร์ชันที่ใช้งานอยู่ใน ปัจจุบันคือ WSDL 2.0

เอกสารดับเบิ้ลยูเอสดีแอลถูกสร้างขึ้นตามมาตรฐานของภาษาเอ็กซ์เอ็มแอล ประกอบด้วยการอธิบายถึงส่วนต่าง ๆ ของเว็บเซอร์วิส รายละเอียดแสดงในตารางที่ 2.1

ตารางที่ 2.1 ส่วนประกอบของเอกสารดับเบิ้ลยูเอสดีแอล

อีลีเมนต์	คำอธิบาย
<portType>	เป็นส่วนที่อธิบายโอเปอเรชันที่เว็บเซอร์วิสมีให้บริการ รวมทั้งข้อความที่เกี่ยวข้อง
<operation>	เป็นส่วนที่ใช้อธิบายเมธอด (Method) ที่ให้บริการ โดยแต่ละเว็บเซอร์วิสสามารถมีได้หลายเมธอด
<message>	เป็นส่วนที่ใช้อธิบายอีลีเมนต์ข้อมูลโอเปอเรชัน โดยแต่ละโอเปอเรชันสามารถมีได้หลายข้อความ เปรียบได้กับพารามิเตอร์ของฟังก์ชันในการเขียนโปรแกรม
<types>	เป็นส่วนที่ใช้อธิบายชนิดของข้อมูลที่ใช้ในเว็บเซอร์วิสเพื่อให้เป็นมาตรฐานกลาง เอกสารดับเบิ้ลยูเอสดีแอล จึงใช้สกีมาของภาษาเอ็กซ์เอ็มแอลอธิบายชนิดของข้อมูล
<binding>	เป็นส่วนที่ใช้อธิบายรูปแบบข้อความ (Message Format) และรายละเอียดของโพรโทคอลในแต่ละพอร์ต
<service>	เป็นเซอร์วิสที่เว็บเซอร์วิสมีไว้ให้บริการ โดยแต่ละเว็บเซอร์วิสสามารถมีได้มากกว่าหนึ่งเซอร์วิส

โครงสร้างของเอกสารดับเบิ้ลยูเอสดีแอล เป็นไปตามแบบมาตรฐานของภาษาเอ็กซ์เอ็มแอล ประกอบด้วยอีลีเมนต์ต่าง ๆ ตัวอย่างโครงสร้างของเอกสารดับเบิ้ลยูเอสดีแอล แสดงในรูปที่ 2.8

```
<wsdl:definitions name="nmtoken"? targetNamespace="uri">
  <import namespace="uri" location="uri"/> *
   <wsdl:documentation .... /> ?
   <wsdl:types> ?
     <wsdl:documentation .... /> ? <xsd:schema .... /> *
   </wsdl:types>
   <wsdl:message name="ncname"> *
     <wsdl:documentation .... /> ? <part name="ncname" element="qname"? type="qname"?/> *
   </wsdl:message>
   <wsdl:portType name="ncname"> *
     <wsdl:documentation .... /> ?
     <wsdl:operation name="ncname"> *
       <wsdl:documentation .... /> ?
       <wsdl:input message="qname"> ? <wsdl:documentation .... /> ? </wsdl:input>
       <wsdl:output message="qname"> ? <wsdl:documentation .... /> ? </wsdl:output>
     </wsdl:operation>
   </wsdl:portType>
   <wsdl:serviceType name="ncname"> *
     <wsdl:portType name="qname"/> +
   </wsdl:serviceType>
   <wsdl:binding name="ncname" type="qname"> *
     <wsdl:documentation .... /> ? <-- binding details --> *
     <wsdl:operation name="ncname"> *
       <wsdl:documentation .... /> ? <-- binding details --> *
       <wsdl:input> ? <wsdl:documentation .... /> ? <-- binding details --> </wsdl:input>
       <wsdl:output> ? <wsdl:documentation .... /> ? <-- binding details --> * </wsdl:output>
     </wsdl:operation>
   </wsdl:binding>
   <wsdl:service name="ncname" serviceType="qname">          *
     <wsdl:documentation .... /> ?
     <wsdl:port name="ncname" binding="qname"> *
       <wsdl:documentation .... /> ? <-- address details -->
     </wsdl:port>
   </wsdl:service>
 </wsdl:definitions>
```

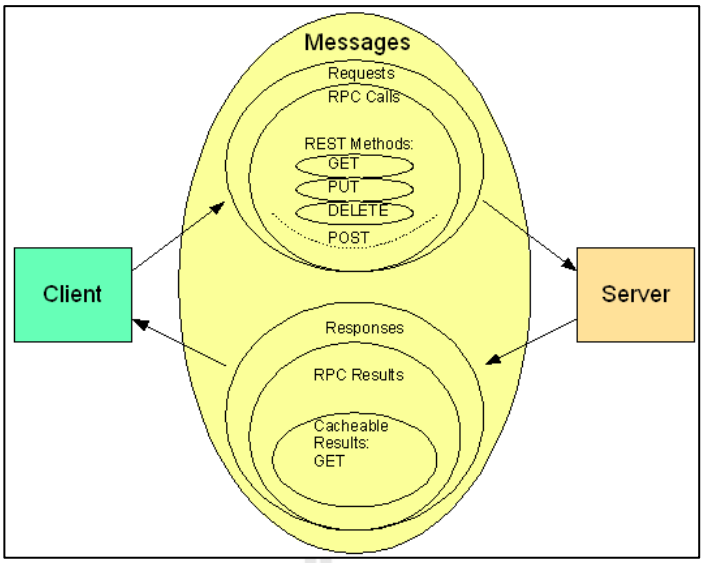
รูปที่ 2.8 ตัวอย่างโครงสร้างของเอกสารดับเบิ้ลยูเอสดีแอล (WSDL)

4) ยูดีดีไอ (Universal Description, Discovery and Integration: UDDI) (UDDI-OASIS Standard, www, 2004)

ยูดีดีไอเป็นไดเรกทอรีที่เก็บรวบรวมเว็บเซอร์วิสและรายละเอียดต่าง ๆ รวมถึงไฟล์หรือที่อยู่ของดับเบิ้ลยูเอสดีแอล ที่ผู้ให้บริการได้ลงทะเบียนไว้เพื่อประกาศบริการของตนเองสู่สาธารณะ และผู้ใช้บริการสามารถค้นหาบริการที่ต้องการใช้งานจากยูดีดีไอได้ ข้อมูลในยูดีดีไอจะประกอบไปด้วยรายละเอียดที่เกี่ยวกับองค์กร (businessEntity) รายละเอียดเกี่ยวกับเซอร์วิส (businessService) รายละเอียดเกี่ยวกับการติดต่อ (bindingTemplate) ที่อยู่หรือยูอาร์แอล (Uniform Resource Locator: URL) สำหรับการเรียกใช้เซอร์วิส (accessPoint) และข้อมูลอ้างอิงไปยัง WSDL (TModelInstanceInfo) นอกจากนี้ยังสามารถแบ่ง Registry ได้เป็นสองประเภทคือ public registry ซึ่งเป็น registry ที่เปิดให้ใช้ทั่วไปทั้งภายใน และภายนอกองค์กร กับ private registry ซึ่งเป็น registry ที่เปิดให้ใช้เฉพาะภายในองค์กร

5) เรสท์ (REpresentational State Transfer: REST)

เรสท์เป็นวิธีการใช้งานเว็บเซอร์วิสรูปแบบหนึ่ง โดยเรียกใช้ผ่านทางโพรโทคอลเอชทีทีพี (HTTP) ด้วยวิธี GET POST PUT และ DELETE และส่งข้อมูลออกมาในรูปของเอ็กซ์เอ็มแอล ผู้เรียกใช้เว็บเซอร์วิสแบบเรสท์เพียงแค่ทราบยูอาร์แอล (Uniform Resource Locator: URL) ของเว็บเซอร์วิส และการอ่านข้อมูลเอ็กซ์เอ็มแอล ก็จะดึงข้อมูลที่ตนเองต้องการได้ ในปัจจุบันนี้มีผู้ให้บริการเว็บเซอร์วิสที่รองรับ RESTful Web Service หรืออีกชื่อหนึ่งคือ RESTful Web API กันเป็นส่วนใหญ่แล้ว ด้วยโครงสร้างของเรสท์ (REST) ที่เรียบง่ายจึงทำให้เรสท์ (REST) มีขนาดเล็ก มีแบนด์วิธที่ต่ำ และทำให้การรับส่งข้อมูลรวดเร็วขึ้น ปัจจุบันเว็บเซอร์วิสแบบเรสท์ (REST) ได้รับความนิยมในการนำมาใช้สร้างเว็บเซอร์วิสมากกว่าแบบโซพ (SOAP) ด้วยความเร็วในการประมวลผลที่รวดเร็วกว่าและอาศัยเทคโนโลยีของเว็บเป็นหลัก ง่ายต่อการเรียกใช้เซอร์วิสโดยสามารถเรียกใช้งานเซอร์วิสผ่านทางยูอาร์แอล มีการทำงานโดยใช้เมธอด GET POST PUT และ DELETE ผ่านโพรโทคอลเอชทีทีพี (HTTP) ดังรูปที่ 2.9 และใช้ภาษาเอ็กซ์เอ็มแอลเป็นสื่อกลางในการแลกเปลี่ยนข้อมูล ซึ่งเรสท์ (REST) สามารถรองรับข้อมูลในการรับส่งได้มากกว่าเว็บเซอร์วิสแบบโซพ (SOAP) (Hatem Haman, 2010; Jan Newmarch, 2009) การอธิบายเซอร์วิสของเรสท์ใช้ภาษา WADL (Web Application Description Language) และการแลกเปลี่ยนข้อมูลของเรสท์สามารถใช้ได้หลายรูปแบบ เช่น XML หรือ JSON



รูปที่ 2.9 รูปแบบโครงสร้างการทำงานของเรสท์ (REpresentational State Transfer: REST)

การนำแนวคิดของเรสท์ (REST) มาใช้งานเว็บเซอร์วิสก็เปรียบเหมือนกับการทำเว็บไซต์แบบปกติทั่วไป โดยฝั่งผู้ใช้ (Client) แทนที่จะเป็นโปรแกรมเว็บบราวเซอร์ ก็สามารถใช้โปรแกรมใดก็ได้ เช่น โปรแกรมระบบคลังสินค้าส่งคำร้องมาที่เครื่องแม่ข่าย (Server) เครื่องแม่ข่ายก็ประมวลผลตามปกติแล้วส่งกลับไปในรูปแบบที่สะดวกกับผู้ใช้ซึ่งส่วนใหญ่แล้วเป็นเอ็กซ์เอ็มแอล เมื่อโปรแกรมฝั่งผู้ใช้ได้รับเอ็กซ์เอ็มแอลกลับมาก็สามารถนำไปประมวลผลต่อได้ทันที ดังนั้นจะเห็นว่าการสร้างเว็บเซอร์วิสแบบเรสท์ (REST) ได้รับความนิยมมากขึ้นเรื่อย ๆ เนื่องจากพัฒนาได้ง่ายทั้งฝั่งผู้ใช้และฝั่งแม่ข่าย ประกอบกับความล้มเหลวของแนวคิดที่จะทำไดเรกทอรีกลางหรือยูดีดีไอ (UDDI) ของโซพ (SOAP) ซึ่งในอดีตมีอยู่ 3 บริษัท ได้แก่ ไอบีเอ็ม (IBM) ไมโครซอฟต์ (Microsoft) และแซพ (SAP) แต่ปัจจุบันได้ยกเลิกไดเรกทอรีกลางไปหมดแล้ว ทำให้ขณะนี้เว็บเซอร์วิสในรูปแบบของเรสท์ (REST) ได้กลายเป็นมาตรฐานไปแล้ว (De Facto Standard)

ผู้ให้บริการเว็บเซอร์วิสหลายรายในปัจจุบันมีการปรับเปลี่ยนและมุ่งเน้นการให้บริการข้อมูลที่เป็นเว็บเซอร์วิสแบบเรสท์มากขึ้น เนื่องจากสามารถให้บริการผู้ใช้ได้ง่ายและใช้เทคโนโลยีที่ไม่ซับซ้อน เพียงแค่ผู้ใช้บริการมีโปรแกรมเว็บบราวเซอร์และศึกษาคู่มือการใช้งานผ่านทางเว็บไซต์ของผู้ให้บริการก็สามารถเรียกใช้บริการได้ ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงเก็บรวบรวมข้อมูลเว็บเซอร์วิสที่เป็นแบบเรสท์เป็นส่วนใหญ่ เพื่อนำมาสร้างเป็นฐานความรู้ในการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว

2.3 การค้นหาข้อมูลเชิงความหมาย

ปัจจุบันการสืบค้นข้อมูลในอินเทอร์เน็ตมีความสำคัญมากและมีบทบาทสำคัญในชีวิตประจำวันของมนุษย์ สังเกตจากการเข้าใช้บริการเว็บไซต์ที่ใช้ค้นหาข้อมูลอย่างกูเกิ้ล (Google) ที่ผู้ใช้อินเทอร์เน็ตทุกคนปฏิเสธไม่ได้ว่าไม่เคยเข้าไป แต่อย่างไรก็ตามเนื่องจากมีข้อมูลจำนวนมากบนอินเทอร์เน็ต อาจทำให้โปรแกรมค้นหา (Search Engine) ไม่สามารถค้นหาเอกสารข้อมูลที่ผู้ใช้ต้องการได้อย่างถูกต้องแม่นยำ ทำให้นักวิจัยทางคอมพิวเตอร์คิดค้นวิธีในการเพิ่มประสิทธิภาพในการค้นหาข้อมูลให้มีประสิทธิภาพมากขึ้น วิธีการนี้เรียกว่า การค้นคืนข้อมูลสารสนเทศ (Information Retrieval) (Manning, Raghavan and Schutze, 2008)

โดยทั่วไปในกระบวนการพื้นฐานของโปรแกรมค้นหาข้อมูลในการค้นคืนข้อมูลสารสนเทศประกอบไปด้วยวิธีการดังนี้

1) การทำดัชนี

เป็นวิธีการสร้างหรือการกำหนดให้ข้อมูลนั้นทำหน้าที่เป็นตัวแทนของเอกสารต่าง ๆ ในฐานข้อมูล ซึ่งโดยปกติแล้วโปรแกรมค้นหาจะไม่ได้ทำการค้นหาจากตัวเอกสารข้อมูลนั้นโดยตรง แต่จะทำการค้นหาข้อมูลที่เป็นตัวแทนเอกสารที่เรียกว่าดัชนี (Index) ซึ่งการทำระบบดัชนีจะช่วยโปรแกรมค้นหาสามารถค้นหาข้อมูลได้รวดเร็วขึ้น กระบวนการทำดัชนีที่เป็นที่รู้จักกันอย่างแพร่หลายคือ วิธีการทำดัชนีแบบอินเวิร์ท (Inverted Indexing) ซึ่งมีขั้นตอนการทำงานที่สำคัญได้แก่ การรวบรวมเอกสารข้อมูลที่ต้องการทำดัชนี การตัดคำ (Tokenizing) การหารากศัพท์เพื่อเปลี่ยนรูปของคำศัพท์ให้อยู่ในรูปแบบดั้งเดิม (Stemming Words) การตัดคำที่ไม่มีความหมายหรือการตัดคำหยุด (Stop Words) และจากนั้นทำการจัดทำดัชนีแบบอินเวิร์ท ซึ่งวิธีการจัดทำดัชนีแบบอินเวิร์ทนี้ถือเป็นวิธีการที่มีประสิทธิภาพและเหมาะสำหรับการทำดัชนีเอกสารที่เป็นตัวอักษร

2) การเปรียบเทียบ

คือการนำเอาคำค้นหรือคำสำคัญมาเปรียบเทียบกับคำในดัชนี วิธีการเดิมของการเปรียบเทียบคือการเปรียบเทียบแบบอักขระ (String Matching) ที่อยู่ในข้อคำถามจากผู้ใช้กับสายอักขระที่อยู่ในดัชนี ถ้าเปรียบเทียบแล้วตรงกัน โปรแกรมสืบค้นจะถือว่าเอกสารที่มีสายอักขระนั้นปรากฏอยู่เป็นเอกสารที่ตรงกับที่ผู้ใช้ต้องการ แต่วิธีการนี้ถือว่าไม่มีประสิทธิภาพเพราะเอกสารบางอย่างอาจมีความเกี่ยวข้องกับข้อคำถามจากผู้ใช้ เพียงแต่ไม่มีคำที่เปรียบเทียบแล้วตรงกันจึงไม่ถูกค้นพบ ดังนั้นจึงมีวิธีการใหม่เพื่อแก้ปัญหานี้เรียกว่า การค้นหาเชิงความหมาย (Semantic Search)

3) การเรียงลำดับผลลัพธ์

เนื่องจากผลลัพธ์ที่ได้จากการค้นหาอาจมีจำนวนมาก จึงต้องมีวิธีจัดการเรียงลำดับผลลัพธ์เพื่อให้ผู้ใช้ได้ข้อมูลที่ตรงกับที่ผู้ใช้ต้องการจริง ๆ โดยสามารถคำนวณได้จากการวัดความ

คล้ายคลึงกันระหว่างคำสำคัญหรือข้อคำถามจากผู้ใช้และเอกสารผลลัพธ์ ซึ่งวิธีที่นิยมสำหรับการคำนวณคือการวัดความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) ซึ่งเป็นวิธีการหาความคล้ายคลึงกันจากค่าความต่างของมุมของข้อมูลสองมุมที่เกิดขึ้นบนพื้นที่เวกเตอร์ วิธีการนี้เป็นวิธีการที่มีประสิทธิภาพและถูกนำมาประยุกต์ใช้กับศาสตร์ทางด้านการค้นคืนสารสนเทศอย่างแพร่หลาย (Salton, 1988; Sidorov et al., 2014)

ถึงแม้ว่าการค้นคืนข้อมูลสารสนเทศจะทำให้การค้นหาข้อมูลมีประสิทธิภาพมากขึ้น แต่ปัญหาที่สำคัญยังพบว่าการค้นหาข้อมูลในโปรแกรมค้นหาส่วนใหญ่จะพิจารณาโดยการดูว่ามีคำสำคัญในข้อคำถาม (Query) ปรากฏอยู่ในเอกสารที่ค้นหาหรือไม่ถ้าไม่มีระบบจะถือว่าเอกสารนั้นไม่เกี่ยวข้อง แต่ความเป็นจริงแล้วเอกสารบางเอกสารอาจจะเกี่ยวข้องกับคำสำคัญจากผู้ใช้ ถึงแม้ว่าจะไม่มีคำสำคัญปรากฏร่วมกันเลยก็ตาม ความสัมพันธ์ในลักษณะนี้เรียกว่า ความสัมพันธ์เชิงความหมาย ดังนั้นแนวคิดที่สำคัญในการพัฒนารูปแบบในการค้นหาข้อมูล คือการค้นหาข้อมูลโดยพิจารณาจากแนวคิด (Concept) หรือความหมายของข้อคำถาม (Query) หรือที่เรียกว่าการค้นหาข้อมูลเชิงความหมาย เนื่องจากมนุษย์มีระบบความคิดและการรับรู้ถึงความหมายของสิ่งต่าง ๆ ได้ง่ายขึ้น ถ้ามีการจัดการข้อมูลในลักษณะของการเชื่อมโยงความสัมพันธ์ของสิ่งต่าง ๆ แสดงในรูปแบบของโครงสร้างลำดับชั้น ดังนั้นจึงเกิดแนวความคิดการทำดัชนีในรูปแบบของโครงสร้างลำดับชั้นที่เรียกว่าออนโทโลยี ซึ่งเป็นลักษณะคล้ายกับโครงสร้างต้นไม้ โครงสร้างนี้จะทำให้กระบวนการค้นหาข้อมูลเชิงความหมายมีประสิทธิภาพมากขึ้น ซึ่งผู้วิจัยได้อธิบายเรื่องของออนโทโลยีและเทคโนโลยีที่เกี่ยวข้องไว้ในหัวข้อถัดไป

กูฮา เมคคูลและมิลเลอร์ (Guha, McCool and Miller, 2003) อธิบายความหมายของคำว่าการสืบค้นเชิงความหมายไว้ว่าเป็นโปรแกรมประยุกต์สำหรับค้นหาข้อมูลของเว็บเชิงความหมาย (Semantic Web) ที่อาศัยพื้นฐานของการค้นคืนสารสนเทศเพื่อทำการค้นหาทรัพยากรสารสนเทศที่มีความเกี่ยวข้องหรือมีความสัมพันธ์เชิงความหมาย ซึ่งทำให้ได้ผลลัพธ์ที่สามารถเพิ่มประสิทธิภาพและปรับปรุงขั้นตอนวิธีการค้นคืนสารสนเทศแบบดั้งเดิมได้

ไกรศักดิ์ เกสร (2554) ได้สรุปเกี่ยวกับการค้นหาข้อมูลเชิงความหมายไว้ว่า องค์ประกอบหลัก ๆ ยังคงใช้องค์ประกอบของวิธีการค้นหาข้อมูลแบบดั้งเดิมและใช้กระบวนการเดิม เช่น การทำดัชนี การเปรียบเทียบ และการเรียงลำดับข้อมูลผลลัพธ์ แต่เพื่อให้ระบบสามารถรองรับการค้นหาข้อมูลเชิงความหมายได้ เทคนิคที่นำมาใช้บางส่วนจึงต้องปรับเปลี่ยน เช่น การทำดัชนีที่จะสนับสนุนการค้นหาข้อมูลเชิงความหมาย ได้แก่ วิธีการ Latent Semantic Indexing (LSI) และวิธีการออนโทโลยี (Ontology)

ดังนั้นผู้วิจัยจึงสรุปได้ว่า การค้นหาข้อมูลเชิงความหมาย หมายถึง กระบวนการในการค้นหาข้อมูลที่พิจารณาว่าข้อมูลที่ต้องการค้นหานั้นมีความสัมพันธ์กันกับคำสำคัญหรือคำค้นหรือไม่ โดยการนำคำค้นจากผู้ใช้ไปกำหนดกลุ่มคำสำคัญอื่น ๆ ที่เกี่ยวข้องกันเชิงความหมายกับคำค้นดังกล่าว จากนั้นนำกลุ่มคำสำคัญที่ได้ไปสืบค้นทรัพยากรสารสนเทศที่อยู่ในฐานความรู้ที่มีโครงสร้างในการจัดเก็บข้อมูลแบบลำดับชั้นและมีการเชื่อมโยงข้อมูลที่มีความเกี่ยวข้องกัน ซึ่งกระบวนการที่สำคัญของการค้นหาข้อมูลเชิงความหมาย ได้แก่ การทำดัชนีโดยวิธีการออนโทโลยี การเปรียบเทียบ และการเรียงลำดับข้อมูลผลลัพธ์

2.4 ออนโทโลยีและเทคโนโลยีที่เกี่ยวข้อง

2.4.1 ทฤษฎีพื้นฐานเกี่ยวกับออนโทโลยี

ออนโทโลยีเป็นแนวคิดที่ปรากฏอยู่ในหลายสาขาวิชาซึ่งแต่ละสาขาวิชาก็ได้ให้ความหมายของออนโทโลยีที่แตกต่างกันไป ในทางคอมพิวเตอร์เองก็มีผู้ให้ความหมายของออนโทโลยีไว้หลากหลาย แต่พอจะสรุปได้ว่าออนโทโลยีคือข้อกำหนดเกี่ยวกับแนวคิด โดยแนวความคิดของออนโทโลยีคือการพยายามที่จะบรรยายแนวคิดของโดเมนหรือขอบเขตความสนใจด้านใดด้านหนึ่งในรูปของสิ่งต่าง ๆ ที่อยู่ภายในโดเมนนั้น สามารถจัดหมวดหมู่ของข้อมูลและเชื่อมโยงความสัมพันธ์ให้กับสิ่งเหล่านั้น ซึ่งสามารถแสดงออกมาในรูปของสัญลักษณ์ เช่น คลาส (Class) อินสแตนซ์ (Instance) ความสัมพันธ์ (Relationship) คุณสมบัติ (Property) และกฎ (Rule) นอกจากนี้ออนโทโลยีที่ถูกสร้างขึ้นมาจะมีความสามารถในการใช้ข้อมูลร่วมกัน (Share) สามารถนำข้อมูลกลับมาใช้ใหม่ได้ (Reuse) และมีความสามารถในการถ่ายทอดคุณสมบัติ (Inheritance) (Chandrasekaran, Josephson and Benjamins, 1999) การนำออนโทโลยีมาใช้งานจึงเป็นทางเลือกหนึ่งในการแชร์ข้อมูล และแยกองค์ความรู้ออกจากฐานข้อมูล โดยใช้ภาษาสำหรับแสดงความรู้ซึ่งมีความชัดเจนและเที่ยงตรงมากกว่าการอธิบายโดยใช้ภาษาธรรมชาติที่ใช้คำศัพท์มาเชื่อมต่อกันเป็นประโยคเพื่อบรรยายถึงสิ่งของในแง่มุมต่าง ๆ ทั้งนี้การใช้ระบบสัญลักษณ์จะช่วยสื่อความหมาย (Semantics) ให้ซอฟต์แวร์หรือคอมพิวเตอร์เข้าใจสามารถให้ความหมายและนำไปใช้ประโยชน์ได้ โดยเฉพาะการค้นหาสิ่งที่ต้องการจากอินเทอร์เน็ตซึ่งเปรียบเสมือนคลังข้อมูลขนาดใหญ่ หากเราสามารถค้นหาข้อมูลได้จากความหมายไม่ใช่จากการค้นหาจากคำเหมือน ก็จะช่วยให้เราได้ข้อมูลผลลัพธ์ที่ตรงตามความต้องการและสะดวกรวดเร็วยิ่งขึ้น ซึ่งตรงกับแนวความคิดของเว็บเชิงความหมาย (Semantic Web)

ปัจจุบันมีโปรแกรมหรือเครื่องมือสำหรับสนับสนุนการพัฒนาออนโทโลยี (Ontology Editor) ที่ได้รับความนิยม เช่น โปรแกรมโพรเทเจ (Protégé) ซึ่งเป็นโปรแกรมแบบโอเพนซอร์ส ผู้ใช้

สามารถสร้างออนโทโลยีได้ง่ายเพราะมีส่วนติดต่อผู้ใช้ที่เป็นแบบกราฟิก พัฒนาโดยมหาวิทยาลัยสแตนฟอร์ด (Musen, 2015) และโปรแกรมโฮโซ (Hozo) พัฒนาโดยมหาวิทยาลัยโอซากา (Mizoguchi, 2007) เป็นต้น โดยเครื่องมือเหล่านี้เป็นเครื่องมือสนับสนุนกระบวนการวิศวกรรมความรู้ที่ช่วยให้ผู้ใช้ที่เป็นวิศวกรความรู้หรือผู้เชี่ยวชาญเฉพาะสาขา สามารถถ่ายทอดและจัดเก็บองค์ความรู้ในรูปแบบของออนโทโลยีได้สะดวกและง่ายดายยิ่งขึ้น

งานวิจัยนี้ได้ประยุกต์ใช้ออนโทโลยีจาก WSMO (Web Service Modeling Ontology) ของดับเบิ้ลยูทรีซี (W3C) ซึ่งเป็นออนโทโลยีของเว็บเซอร์วิสมีมาตรฐานและได้รับการยอมรับ (W3C-WSMO, www, 2005) โดยนำออนโทโลยีนี้มาปรับปรุงให้เข้ากับบริบทด้านการท่องเที่ยว มีการเพิ่มคลาสออนโทโลยีใหม่ มีการกำหนดความสัมพันธ์เพิ่มเติม มีการกำหนดคุณสมบัติต่าง ๆ มีการเพิ่มอินสแตนซ์ และมีการสร้างคลังคำ (Word Repository) โดยใช้สคอส (Simple Knowledge Organization System: SKOS) (SKOS, www, 2009)

2.4.2 องค์ประกอบของออนโทโลยี

1) แนวคิดหรือคลาส (Concept/Class) หมายถึง แนวคิดในขอบเขตความรู้เรื่องใดเรื่องหนึ่ง ซึ่งสามารถเป็นอะไรก็ได้ที่เกี่ยวกับสิ่งที่ถูกกล่างถึง และสามารถอธิบายรายละเอียดได้ เช่น งาน (Task) หน้าที่ (Function) การกระทำ (Action) กลยุทธ์ (Strategy) กระบวนการอย่างมีเหตุผล (Reasoning Precess)

2) คุณสมบัติ (Property) หมายถึง คุณสมบัติต่าง ๆ ที่เกี่ยวข้องสัมพันธ์กับแนวคิดหรือคลาส เพื่อนำมาใช้อธิบายและกำหนดถึงคุณสมบัติต่าง ๆ ให้กับแนวคิด/คลาส

3) ความสัมพันธ์ (Relation) เป็นการนำเสนอประเภทของการปฏิสัมพันธ์ระหว่างแนวคิดกับแนวคิด โดยมีการกำหนดความสัมพันธ์ในลักษณะต่าง ๆ ดังนี้

- ความสัมพันธ์ตามลำดับชั้นแบบแทกโซโนมี (Hierarchical/Taxonomy Relationship) คือความสัมพันธ์ของแนวคิดที่อยู่ตามลำดับชั้น ระหว่างคลาสหลักและคลาสย่อย เช่น ความสัมพันธ์แบบ Subclass-of หรือ Is-a hierarchy
- ความสัมพันธ์แบบ Mereology (Assoication/Mereology Ralationship) คือความสัมพันธ์ของแนวคิดที่ไม่อยู่ตามลำดับชั้น เช่น ความสัมพันธ์แบบเป็นส่วนประกอบ Part-of

4) ตรรกะการแปลงความสัมพันธ์ เป็นตรรกะการแปลงความสัมพันธ์ระหว่างแนวคิดกับคุณสมบัติ หรือแนวคิดกับแนวคิด หรือการตรวจสอบคุณสมบัติของคุณสมบัติ เพื่อการแปลงความหมายที่ถูกต้อง

5) ตัวอย่างข้อมูลหรือคำศัพท์ (Instance) เป็นคำศัพท์ที่มีการกำหนดนิยามความหมายร่วมกันทั้งหมดในออนโทโลยี

2.4.3 กระบวนการในการพัฒนาออนโทโลยี

ประกอบไปด้วย 7 ขั้นตอน (Noy and McGuinness, 2000)

1) ระบุขอบเขตและวัตถุประสงค์ของการพัฒนาออนโทโลยี (Determine Domain and Scope)

1.1) ขอบเขตขององค์ความรู้ครอบคลุมเรื่องใดบ้าง

1.2) วัตถุประสงค์ของการนำออนโทโลยีไปใช้ประโยชน์เพื่ออะไร

1.3) การตั้งคำถาม (Competency Question) ที่ต้องการให้ออนโทโลยีที่จะสร้างขึ้นสามารถตอบคำถามเหล่านั้นได้

2) พิจารณาการนำออนโทโลยีเดิมที่มีอยู่กลับมาใช้ใหม่ (Consider Reuse)

3) พิจารณาและแจกแจงประเด็น/หัวข้อทั้งหมดที่ต้องการสื่อเกี่ยวกับออนโทโลยี (Enumerate Important Terms)

3.1) พิจารณาประเด็น/หัวข้อ (Terms) ทั้งหมดที่ต้องการสื่อ

3.2) พิจารณาคุณสมบัติ (Properties) ของประเด็น/หัวข้อนั้น

3.3) ความหมายของประเด็น/หัวข้อต้องการสื่อคืออะไร

4) กำหนดคลาสและลำดับของคลาส (Define Class and Hierarchy) ที่มีอยู่ในโดเมนสำหรับออนโทโลยี

5) กำหนดคุณสมบัติของคลาส (Define Properties of Classes) โดยการบรรยายส่วนประกอบของตัวอย่างข้อมูล (Attributes of Instance – Slots)

6) กำหนดมุมมองและข้อจำกัดของคุณสมบัติ (Define Constraints) โดยอธิบายและกำหนดค่าของข้อมูลที่เป็นไปได้

7) สร้างตัวอย่างของข้อมูลในคลาส (Create Instances of a Class)

2.4.4 งานวิจัยที่เกี่ยวข้องกับออนโทโลยี

ซูกูมาเรน และสโตเรย์ (Sugumaran and Storey, 2002) ได้นำเสนอวิธีการในการพัฒนาออนโทโลยี โดยแบ่งเป็น 4 ขั้นตอน คือ 1) กำหนดคำศัพท์ทั่วไปสำหรับสร้างออนโทโลยี รวมทั้งกำหนดคำเหมือน 2) กำหนดความสัมพันธ์ระหว่างคำศัพท์เหล่านั้น ได้แก่ ความสัมพันธ์ทั่วไป (Is-a) ความสัมพันธ์แบบคำเหมือน และความสัมพันธ์แบบคำที่เกี่ยวข้องกัน 3) กำหนดเงื่อนไขของความสัมพันธ์เบื้องต้น เช่น ความสัมพันธ์ที่จะเกิดขึ้นได้หรือไม่สามารถเกิดขึ้นได้ ถ้ามีความสัมพันธ์อื่นเกิดขึ้นก่อน และ 4) กำหนดเงื่อนไขระดับสูง เช่น โดเมนและเรนจ์ของความสัมพันธ์ของข้อมูล

สเตบและคณะ (Staab et al., 2001) ได้นำเสนอวิธีการที่ใกล้เคียงกับนอยและแมคกินเนสส์ (Noy and McGuinness, 2000) แต่ลดจำนวนขั้นตอนและเพิ่มวัฏจักรการพัฒนาไปในวิธีการ

ออกแบบและพัฒนาออนโทโลยี โดยงานวิจัยนี้แบ่งขั้นตอนออกเป็น 5 ขั้นตอน คือ 1) ศึกษาความเป็นไปได้ในการพัฒนาออนโทโลยี โดยเป็นการวิเคราะห์ปัญหา และวิธีการนำออนโทโลยีไปใช้งาน 2) กำหนดขอบเขต (Domain) และจัดทำเอกสารข้อกำหนดความต้องการของออนโทโลยี ซึ่งมีรายละเอียดต่าง ๆ เช่น วัตถุประสงค์ของออนโทโลยี ขอบเขตและโดเมน แหล่งความรู้ ผู้ใช้งาน คำถามพื้นฐาน สำหรับกำหนดความสามารถในการตอบคำถามของออนโทโลยี เป็นต้น 3) การรวบรวมความรู้และสร้างออนโทโลยี โดยขั้นตอนนี้จะกำหนดคลาส ความสัมพันธ์ และเงื่อนไขในออนโทโลยี 4) ขั้นตอนการทดสอบออนโทโลยี โดยเริ่มแรกผู้พัฒนาจะต้องตรวจสอบตามเอกสารข้อกำหนดความต้องการของออนโทโลยีที่ได้กำหนดไว้ และตรวจสอบว่าออนโทโลยีที่ออกแบบไว้สามารถตอบคำถามพื้นฐานที่กำหนดไว้ได้หรือไม่ และให้ผู้ใช้ตรวจสอบความพึงพอใจของโครงสร้างออนโทโลยี โดยจากขั้นตอนนี้ผู้พัฒนาสามารถกลับไปแก้ไขในขั้นตอนที่ 3 ได้ และ 5) การปรับปรุงและบำรุงรักษาออนโทโลยี เนื่องจากข้อมูลความรู้อาจมีการเปลี่ยนแปลงได้ตลอดเวลา ดังนั้นในขั้นตอนนี้จึงเป็นการแก้ไขและปรับปรุงโครงสร้างหรือขอบเขต

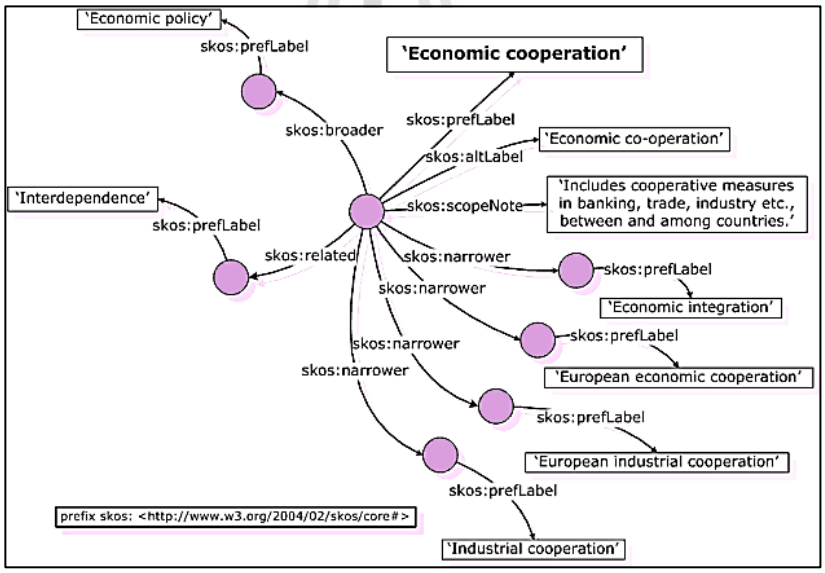
อนุวัฒน์ โยวบุตร และฉัตรตระกูล สมบัติ ธีระ (2552) ได้นำเสนอวิธีการค้นหาข้อมูลด้วยการประยุกต์เว็บเชิงความหมาย โดยนำเทคโนโลยีเว็บเชิงความหมายมาประยุกต์กับการค้นหาข้อมูล โดยเฉพาะเทคนิคการเชื่อมโยงความสัมพันธ์ของข้อมูล ประเด็นหลักที่ทำให้เกิดการนำเว็บเชิงความหมายมาประยุกต์ คือ ข้อมูลที่สืบค้นมาได้ไม่มีประสิทธิภาพเพียงพอ และไม่สะดวกในการที่จะนำไปใช้ต่อ เพราะการค้นหาด้วยคำสำคัญ เครื่องคอมพิวเตอร์ไม่สามารถเข้าใจถึงความหมายหรือความสัมพันธ์ของคำ ผลของการสืบค้นจะแสดงข้อมูลเฉพาะรายการที่มีคำ ๆ นั้นปรากฏอยู่ซึ่งอาจไม่ตรงกับความหมายที่ต้องการ แนวทางของเว็บเชิงความหมายสามารถแก้ปัญหาดังกล่าวด้วยการทำให้ข้อมูลมีการระบุขอบเขตได้ โดยที่เครื่องคอมพิวเตอร์สามารถเข้าใจองค์ประกอบของข้อมูลซึ่งมีแนวคิดตามทฤษฎีของโดเมน เช่น รูปแบบของการอ้างอิงคลาสแม่ของข้อมูลเรียกว่าออนโทโลยี ซึ่งสามารถบอกระดับความสัมพันธ์ของข้อมูลได้จึงมีประโยชน์อย่างมากในการค้นหาข้อมูล เป็นการเชื่อมโยงความสัมพันธ์ของข้อมูลทำให้ได้ผลลัพธ์ที่ถูกต้อง มีประสิทธิภาพและตรงตามความต้องการของผู้ใช้งาน

อรวรรณ อุไรเรืองพันธ์ และสมจิตร อาจอินทร์ (2552) ทำวิจัยเกี่ยวกับการสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยี ซึ่งได้นำแนวคิดการสรุปใจความอัตโนมัติ เพื่อช่วยให้ผู้ใช้สามารถลดเวลาในการคัดเลือกข้อมูลเอกสารอิเล็กทรอนิกส์ที่ต้องการค้นหาได้ถูกต้องและมีประสิทธิภาพมากขึ้น โดยเลือกใช้เทคโนโลยีด้านความคล้ายคลึงกันเชิงความหมาย ที่ช่วยลดปัญหาเรื่องคำกำกวมจากการสรุปใจความแบบเดิมที่ใช้การอ้างอิงคำสำคัญ และวิธีการสรุปใจความสำคัญซึ่งเป็นการถอดความหรือแปลความหมายคำจากต้นฉบับด้วยหลักเชิงความหมาย จากนั้นจัด

หมวดหมู่ของกลุ่มความรู้ที่ได้จากการสรุปใจความสำคัญ โดยพิจารณาลักษณะความสัมพันธ์เชิงความหมายระหว่างคำหลัก โดยอาศัยฐานข้อมูลคำศัพท์เวิร์ดเน็ต (George A. Miller, 1995) และคำนวณค่าน้ำหนัก และเชื่อมโยงสัมพันธ์กันตามทฤษฎีออนโทโลยีผลการวิจัยพบว่าการสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยีจัดกลุ่มคำหลักช่วยให้ได้ผลลัพธ์ในเชิงความหมายที่มีประสิทธิภาพมากขึ้น

2.4.5 สคอส (SKOS)

สคอส (Simple Knowledge Organization System: SKOS) (SKOS, www, 2009) เป็นตัวแบบสำหรับสร้างคลังคำหรือมาตรฐานข้อมูลที่ใช้ร่วมกัน พื้นฐานของสคอสประกอบไปด้วยความสัมพันธ์ดังนี้ คือ คำหลัก (skos:preLabel) คำเหมือน (skos:altLabel) คำที่กว้างกว่า (skos:broader) คำที่แคบกว่า (skos:narrower) และคำที่เกี่ยวข้อง (skos:related) (Fulvio and Laura, 2005) ตัวอย่างของการนิยามถึงความสัมพันธ์ของคำตามแบบโครงสร้างของสคอส แสดงได้ดังรูปที่ 2.10



รูปที่ 2.10 โครงสร้างความสัมพันธ์ของคำศัพท์แบบสคอส (SKOS, www, 2009)

2.4.6 เวิร์ดเน็ต (WordNet)

เวิร์ดเน็ต (George A. Miller, 1995) เป็นฐานข้อมูลคำศัพท์ภาษาอังกฤษขนาดใหญ่ประกอบด้วยคำนาม คำกริยา คำคุณศัพท์ และกริยาวิเศษณ์ มีการจัดกลุ่มเป็นชุดของคำตามกลุ่มความหมาย โดยจะเชื่อมโยงโดยวิธีการของความสัมพันธ์ของแนวคิดและความหมายคำศัพท์ตามลำดับชั้น ความสัมพันธ์ของความหมายในเวิร์ดเน็ตมีดังนี้

1) Synonymy คือ เซ็ตของคำที่มีความหมายคล้ายคลึงกัน (Synsets, Synonym sets) ในการแทนคำที่มีความหมาย

2) Antonymy คือคำที่มีความหมายตรงกันข้ามกันอย่างสมมาตรซึ่งเน้นที่คำประเภทคำคุณศัพท์และคำกริยาวิเศษ

3) Hyponymy และ Hypernymy คือความสัมพันธ์ของคอนเซ็พที่มีลักษณะของ Concept และ Super Concept โดย Hyponym คือคอนเซ็พที่อยู่เหนือคอนเซ็พอื่น ส่วน Hypernym คือคอนเซ็พที่อยู่ภายใต้คอนเซ็พอื่น เช่น Hotel มี Hypernymy คือ building และ building มี Hyponymy คือ Hotel เป็นต้น

4) Meronymy และ Holonymy คือความสัมพันธ์แบบบางส่วนและทุกส่วน เช่น แขน ขา มือ เป็นส่วนหนึ่งของร่างกาย และพวงมาลัย เบรก ล้อ เป็นส่วนหนึ่งของรถ เป็นต้น

5) Troponymy คือคำกริยาเพื่อแสดงอาการหรือการกระทำ ซึ่งใช้สำหรับการแสดงการกระทำของคำนามในลำดับชั้นของ Concept

6) Entailment คือความสัมพันธ์ระหว่างคำกริยาซึ่งกำหนดโดยเวิร์ดเน็ต

2.4.7 ภาษาสปาร์เคิล (SPARQL)

ภาษาสปาร์เคิล (W3C-SPARQL, www, 2013) เป็นภาษามาตรฐานในการใช้สำหรับสืบค้นข้อมูลที่ถูกจัดเก็บในรูปแบบอาร์ดีเอฟ (RDF) หรือโอดับเบิ้ลยูแอล (OWL) (Sbodio et al, 2010) มาแสดงหรือเรียกว่าภาษาสอบถาม (Query Language) โดยอยู่บนพื้นฐานของข้อมูลที่เป็นไปในรูปแบบของกราฟที่ถูกจัดเก็บในรูปแบบของโอดับเบิ้ลยูแอล (OWL) กราฟที่มีลักษณะที่ง่ายที่สุดก็คือเป็นกราฟในรูปแบบของ Triple (Basic Graph Pattern) ที่ประกอบด้วย Subject Predicate และ Object โครงสร้างหลักของภาษาสปาร์เคิลประกอบไปด้วย 3 ส่วนสำคัญ ได้แก่ ส่วนแรก PREFIX คือการกำหนดการอ้างอิงข้อมูลเอกสารโอดับเบิ้ลยูแอล (OWL) โดยประกาศไว้ที่ส่วนหัวของเอกสาร ส่วนที่สอง SELECT ใช้สำหรับระบุตัวแปรที่จะแสดงผล โดยจะต้องมีเครื่องหมาย ? นำหน้าตัวแปร เช่น ?age (อายุ) เป็นต้น ซึ่งเป็น subject ตามรูปแบบใน Basic Graph Pattern และส่วนสุดท้าย WHERE คือเงื่อนไขในการสืบค้นข้อมูล

```
PREFIX ns1: <http://www.w3.org/2001/vcard-rdf/3.0#>
        ns2: <http://sampleVocabulary.org/1.3/People#>
SELECT ?givenName, ?age
FROM <employee.rdf>
WHERE {?x ns1:N ?blank.
          ?blank ns1:Given ?givenName.
          ?x ns2:age ?age.
          FILTER (?age > 20)
 }
```

รูปที่ 2.11 ตัวอย่างภาษาสปาร์เคิล (SPARQL) ในการค้นหาข้อมูลในออนโทโลยี

2.5 การคัดเลือกเว็บเซอร์วิสเชิงความหมาย (Semantic Web Services Selection)

จุดเริ่มต้นของเทคโนโลยีเว็บเกิดขึ้นเมื่อปี ค.ศ. 1994 ทิมเบอร์เนอร์ลี (Tim Berners-Lee) ได้ก่อตั้งองค์กรเว็บไซต์สากล (World Wide Web Consortium: W3C) ขึ้นที่สถาบันเทคโนโลยีแมสซาชูเซตส์ เพื่อสร้างมาตรฐานและข้อเสนอแนะสำหรับใช้เป็นหลักในการพัฒนาและปรับปรุงคุณภาพของเว็บไซต์ หลังจากนั้นในปี ค.ศ. 2004 องค์กร W3C ได้เริ่มโครงการซีแมนติกเว็บ (Semantic Web) หรือที่เรียกว่าเว็บเชิงความหมาย เป็นพัฒนาการของเวิลด์ไวด์เว็บที่เกี่ยวข้องกับข้อมูลและการบริการบนเว็บไซต์ ในการสร้างความเป็นไปได้ที่เว็บไซต์จะสามารถเข้าใจถึงความต้องการของผู้ใช้และเครื่องมือที่ใช้บรรจุลงในสารบัญเว็บไซต์ เพื่อพัฒนาเว็บให้เป็นแหล่งรวมการแลกเปลี่ยนข้อมูลและความรู้ (W3C, www, 2013)

เว็บเชิงความหมายเป็นวิธีการที่จะทำให้คอมพิวเตอร์หรือแอปพลิเคชันสามารถเข้าใจข้อมูลที่สอดคล้องกับความเข้าใจของมนุษย์ สามารถเข้าใจความหมายของคำและแนวความคิดที่รวบรวมเกี่ยวกับคำในขอบเขตของความรู้ที่ต้องการตามที่ผู้พัฒนาได้กำหนดไว้โดยเชื่อมโยงความสัมพันธ์ของข้อมูลในระดับเมทาดาตา (Metadata) แบบมีโครงสร้าง ช่วยทำให้การค้นหาข้อมูลมีความถูกต้องและตรงตามความต้องการของผู้ใช้มากขึ้น เว็บเชิงความหมายจึงเป็นอีกแนวคิดหนึ่งที่ถูกนำไปประยุกต์ใช้ในการพัฒนาระบบการค้นคืนสารสนเทศ (Information Retrieval: IR) ในรูปแบบที่เรียกว่า การสืบค้นเชิงความหมาย (Semantic Search) ซึ่งเป็นวิธีการค้นหาทรัพยากรสารสนเทศที่ต้องการจากคลังสารสนเทศหรือฐานข้อมูลภายใต้เงื่อนไขหรือกระบวนการคัดเลือกที่ผู้ใช้ต้องการ ทำให้สามารถเรียกข้อมูลที่ต้องการออกมาแสดงผลได้อย่างรวดเร็วและครอบคลุมถึงการค้นหาข้อมูลที่มีความเกี่ยวข้องและมีความสัมพันธ์เชิงความหมายด้วย ดังนั้นเพื่อให้การออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสนี้สามารถสืบค้นเว็บเซอร์วิสเชิงความหมายได้อย่างมีประสิทธิภาพ ผู้วิจัยจึงได้ศึกษาแนวคิดและทฤษฎีที่เกี่ยวข้องกับการสืบค้นเชิงความหมาย โดยมีรายละเอียดดังนี้

พนิดา ตันศิริ (2553) ได้กล่าวว่า การสืบค้นข้อมูลเชิงความหมาย เป็นรูปแบบหนึ่งของการประยุกต์ใช้เว็บเชิงความหมาย (Semantic Web) ที่มีความชาญฉลาดและสืบค้นข้อมูลแบบอิงความหมายได้ โดยเว็บเชิงความหมายจะมีข้อมูลที่มีการเชื่อมโยงข้อมูลกันในลักษณะเครือข่ายเชิงความหมาย (Semantic Network) ด้วยเทคโนโลยีเว็บเชิงความหมาย ที่ทำให้ผู้ใช้สามารถสร้างข้อมูล สร้างคำศัพท์ฐานความรู้ และข้อกำหนดในการจัดการข้อมูลได้

ไกรศักดิ์ เกสร (2554) ได้อธิบายถึงความสัมพันธ์ระหว่างการค้นคืนสารสนเทศกับวิธีการสืบค้นเชิงความหมายว่า การสืบค้นเชิงความหมายเป็นส่วนหนึ่งของวิธีการค้นคืนสารสนเทศจากลักษณะสารสนเทศ 2 ประเภทคือ ข้อมูลอักขระ (Textual Information) และข้อมูลมัลติมีเดีย (Multimedia Information) เช่น ภาพ วีดีโอ และเสียง เป็นต้น ซึ่งคำว่าการสืบค้นเชิงความหมายจะหมายถึง วิธีการค้นหาข้อมูลที่นำแนวคิด (Concept) หรือความหมายของข้อคำถาม (Query) มาใช้ในการพิจารณาว่าทรัพยากรสารสนเทศที่ต้องการค้นหานั้น มีความสัมพันธ์เกิดขึ้นร่วมกันกับข้อคำถามที่ต้องการหรือไม่ ถึงแม้ว่าจะไม่มีปรากฏคำค้นที่ตรงกันเลยก็ตาม โดยเรียกความสัมพันธ์ในลักษณะนี้ว่า ความสัมพันธ์เชิงความหมาย

ศุภกฤษฎิ์ นิวัฒนากูล (2556, หน้า 6) ได้นิยามความหมายของคำว่า การสืบค้นเชิงความหมายว่า หมายถึง กระบวนการค้นหาที่มีการนำคำสำคัญไปหาคำสำคัญอื่น ๆ ที่เกี่ยวข้อง แล้วใช้กลุ่มคำสำคัญเหล่านั้นในการสืบค้นความรู้ที่มีการจัดทำอธิบายในฐานความรู้

สมเกียรติ สอนนวล, งามนิจ อาจอินทร์, สมจิตร อาจอินทร์ และ ธีระยุทธ ทองเครือ (2558) ได้ศึกษาถึงการสืบค้นเว็บเซอร์วิสแบบสื่อความหมายโดยใช้ซอฟต์แวร์เอเจนต์ โดยนำเสนอวิธีการเก็บรวบรวมเอกสารอธิบายเว็บเซอร์วิสจากแหล่งต่างๆ ในอินเทอร์เน็ต โดยใช้ตัวแทนการค้นหาหรือซอฟต์แวร์เอเจนต์ ช่วยเก็บรวบรวมข้อมูลคำ อธิบายเว็บเซอร์วิส แล้วนำข้อมูลที่ได้มาวิเคราะห์หาคำสำคัญและให้คะแนนความสำคัญเพื่อจัดกลุ่มคำและเอกสารเว็บเซอร์วิสที่สัมพันธ์กัน โดยอาศัยฐานข้อมูลเวิร์ดเน็ตช่วยในการพิจารณาความคล้ายคลึงกันของคำอย่างสื่อความหมาย ผลลัพธ์ของคำที่มีการจัดกลุ่มกันจะถูกเก็บอยู่ในโครงสร้างออนโทโลยีเพื่อให้ผู้ใช้งานเว็บเซอร์วิสเข้ามาสืบค้นเว็บเซอร์วิสที่ต้องการได้จากระบบอย่างสื่อความหมาย ทำให้ผู้ใช้งานสามารถสืบค้นข้อมูลเว็บเซอร์วิสได้ตรงกับความต้องการ

ดังนั้นผู้วิจัยจึงสรุปได้ว่า การสืบค้นข้อมูลเชิงความหมาย หมายถึง กระบวนการในการค้นหาข้อมูลที่พิจารณาว่าข้อมูลที่ต้องการค้นหานั้นมีความสัมพันธ์กันกับคำสำคัญหรือคำค้น (Keywords) หรือไม่ โดยการนำคำค้นจากผู้ใช้ไปกำหนดกลุ่มคำสำคัญอื่น ๆ ที่เกี่ยวข้องเชิงความหมายกับคำค้นดังกล่าว จากนั้นนำกลุ่มคำสำคัญที่ได้ไปสืบค้นทรัพยากรสารสนเทศที่อยู่ในฐานความรู้

2.6 คุณภาพการให้บริการของเว็บเซอร์วิส

คุณภาพการให้บริการ (Quality of Service: QoS) เป็นเรื่องที่สำคัญสำหรับผู้ใช้บริการเว็บแอปพลิเคชันที่มีกระบวนการส่งข้อมูลสารสนเทศผ่านเครือข่ายอินเทอร์เน็ต เทคโนโลยีเว็บเซอร์วิสเป็นเทคโนโลยีหนึ่งที่ผู้ให้บริการนิยมนำมาใช้ในดำเนินงานทางธุรกิจเพื่อใช้ในการนำเสนอสินค้าหรือบริการ ทำให้จำนวนเว็บเซอร์วิสบนอินเทอร์เน็ตมีจำนวนเพิ่มขึ้นอย่างรวดเร็ว เพราะผู้ให้บริการเว็บเซอร์วิส (Service Provider) มีการเสนอฟังก์ชันการบริการที่หลายหลายมากขึ้น (Li et al, 2007) และในการให้บริการในแต่ละกลุ่มของเว็บเซอร์วิสนั้นมีเว็บเซอร์วิสที่ให้บริการอยู่หลากหลาย ยกตัวอย่างเช่น กลุ่มเว็บเซอร์วิสที่ให้บริการด้านที่พักซึ่งมีเป็นจำนวนมาก ผู้ใช้จะทราบได้อย่างไรว่าเว็บเซอร์วิสใดมีคุณภาพการบริการที่ดีที่สุด ดังนั้นเพื่อให้ผู้ใช้ได้ใช้บริการเว็บเซอร์วิสที่มีคุณภาพและประสิทธิภาพ เรื่องคุณภาพการให้บริการจึงเป็นปัจจัยที่สำคัญในการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวในงานวิจัยนี้

2.6.1 ปัจจัยที่ใช้กำหนดคุณภาพการให้บริการของเว็บเซอร์วิส

จากการศึกษางานวิจัยเกี่ยวกับคุณภาพการให้บริการของเว็บเซอร์วิส พบว่ามีการนำปัจจัยด้านคุณภาพมาใช้เป็นตัวกำหนดถึงคุณภาพการให้บริการที่แตกต่างกันไปตามบริบทและโดเมนของงานวิจัยนั้น ๆ สรุปได้ดังนี้

คาร์โดโซ และคณะ (Cardoso et al., 2004) ได้นำเสนอแบบจำลองคุณภาพการให้บริการของเว็บเซอร์วิสโดยใช้เวิร์กโฟลว์และกระบวนการเว็บเซอร์วิส โดยใช้ปัจจัยคุณภาพ ได้แก่ เวลาตอบสนองของการให้บริการ (Response Time) ความน่าเชื่อถือ (Reliability) และค่าใช้จ่ายในการใช้บริการ (Cost)

ซูซิลาและวาดิเวล (Susila and Vadivel, 2014) ได้ออกแบบและพัฒนาเครื่องมือในการวัดค่าคุณภาพการใช้บริการสำหรับการคัดเลือกเว็บเซอร์วิส โดยใช้ปัจจัยคุณภาพได้แก่ สภาพความพร้อมในการให้บริการ (Availability) ความน่าเชื่อถือ (Reliability) เวลาตอบสนองของการให้บริการ (Response Time) ความสำเร็จในการใช้บริการ (Successability) และความสำเร็จในการนำส่งข้อมูล (Throughput)

มูลูกานาเดม กูพู และชานมูกา (Muruganadam, Gopu and Shanmuga, 2014) ได้นำเสนอแบบจำลองการค้นคืนเว็บเซอร์วิสเชิงความหมายบนพื้นฐานของการประเมินคุณภาพการให้บริการ โดยมีปัจจัยคุณภาพที่เกี่ยวข้องในงานวิจัยนี้ ได้แก่ สภาพความพร้อมในการให้บริการ (Availability) ความน่าเชื่อถือ (Reliability) เวลาตอบสนองของการให้บริการ (Response Time) และความสำเร็จในการนำส่งข้อมูล (Throughput)

เทาและเกวเจย์ (Tao and Kwei-Jay, 2004) ได้ออกแบบและพัฒนาเทคนิควิธีการคัดเลือกเว็บเซอร์วิสโดยใช้เงื่อนไขคุณภาพการบริการ โดยมีปัจจัยคุณภาพที่เกี่ยวข้องในงานวิจัยนี้ ได้แก่ สภาพความพร้อมในการให้บริการ (Availability) ค่าใช้จ่ายในการใช้บริการ (Cost/Price) ความน่าเชื่อถือ (Reliability) และเวลาตอบสนองของการให้บริการ (Response Time)

และงานวิจัยอื่น ๆ ที่เกี่ยวกับคุณภาพการให้บริการของเว็บเซอร์วิส สามารถสรุปได้ดังตารางที่ 2.2

ตารางที่ 2.2 สรุปงานวิจัยเกี่ยวกับปัจจัยด้านคุณภาพการให้บริการ (Quality of Service: QoS)

งานวิจัย / ปัจจัย QoS	Availability	Cost / Price	Latency	Reliability	Response time	Successability	Throughput
Cardoso et al. (2004)	-	✓	-	✓	✓	-	-
Susila and Vadivel (2014)	✓	-	-	✓	✓	✓	✓
Karthiban (2014)	✓	-	✓	✓	✓	✓	✓
Ajao, Deris and Obasa (2013)	✓	-	-	✓	✓	✓	-
Muruganadam, Gopu and Shanmuga (2014)	✓	-	-	✓	✓	-	✓
Rajendran, Balasubramanie and Cherian (2010)	✓	✓	-	-	✓	-	✓
Tao and Kwei-Jay (2004)	✓	✓	-	✓	✓	-	-
Zhang (2014)	✓	-	✓	✓	✓	-	✓
Parhi, Pattanayak and Patra (2014)	✓	✓	-	✓	✓	-	-

2.6.2 ปัจจัยคุณภาพการให้บริการของเว็บเซอร์วิส

จากการทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้องด้านปัจจัยคุณภาพการให้บริการของเว็บเซอร์วิส สำหรับการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวนี้ ผู้วิจัยจะใช้ปัจจัยคุณภาพจำนวน 5 ปัจจัย ซึ่งเป็นปัจจัยที่นักวิจัยส่วนใหญ่ให้ความสำคัญเป็นลำดับต้น ๆ ทางด้านคุณภาพในการให้บริการ มีดังนี้

1) เวลาตอบสนองของเซิร์ฟเวอร์ (Response Time) คือ ระยะเวลาในการประมวลผลของเซิร์ฟเวอร์ที่ให้บริการ โดยเริ่มตั้งแต่ผู้ขอใช้บริการส่งคำร้องขอไปยังเซิร์ฟเวอร์ที่ให้บริการจนกระทั่งได้รับข้อมูลตอบกลับมา สูตรการคำนวณแสดงดังสมการที่ 2.1

$$T = \frac{\sum_{i=1}^{n} (T_r - T_s)}{n} \qquad (2.1)$$

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

โดยที่ T คือ ค่าเวลาตอบสนองโดยเฉลี่ย หน่วยเป็นมิลลิวินาที

T_r คือ เวลาที่ได้รับข้อมูลกลับมา หน่วยเป็นมิลลิวินาที

T_s คือ เวลาเริ่มต้นที่ทำการส่งคำร้องขอ หน่วยเป็นมิลลิวินาที

RT คือ ค่าความน่าจะเป็นของค่าเวลาตอบสนองโดยเฉลี่ย

2) สภาพความพร้อมในการให้บริการ (Availability) คือ สภาพของเว็บเซอร์วิสที่ปรากฏอยู่และผู้ใช้บริการสามารถเข้าใช้บริการได้ในทันที โดยคำนวณจากเวลาตอบสนองในแต่ละครั้งที่ผู้ใช้บริการสามารถเข้าใช้บริการได้ตามช่วงเวลาที่กำหนด วัดจากอัตราส่วนของผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการสามารถเข้าใช้บริการเซอร์วิสได้สำเร็จส่วนด้วยผลรวมเวลาทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสในช่วงเวลาที่กำหนด สูตรการคำนวณแสดงดังสมการที่ 2.2

$$A = \frac{N_{suc}}{N_{all}} \qquad (2.2)$$

โดยที่ A คือ ค่าสภาพความพร้อมในการให้บริการ

N_{suc} คือ ผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการสามารถเข้าใช้บริการเซอร์วิสได้สำเร็จ ในช่วงเวลาที่กำหนด มีหน่วยเป็นมิลลิวินาที

N_{all} คือ ผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิส ในช่วงเวลาที่กำหนด มีหน่วยเป็นมิลลิวินาที

มหาวิทยาลัยเทคโนโลยีสุรนารี

3) ความเสถียรภาพหรือความน่าเชื่อถือของระบบ (Reliability) คือ ระดับความสามารถในการให้บริการได้ของเว็บเซอร์วิส แสดงได้ด้วยอัตราส่วนของจำนวนครั้งของการเรียกใช้บริการเซอร์วิสที่ประสบความสำเร็จ คำนวณได้โดยวัดจากอัตราส่วนของจำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเซอร์วิสได้สำเร็จส่วนด้วยจำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเซอร์วิส ในช่วงเวลาที่กำหนด สูตรการคำนวณแสดงดังสมการที่ 2.3

$$R = \frac{T_{suc}}{T_{all}} \tag{2.3}$$

โดยที่ R คือ ค่าความเสถียรของระบบ

T_{suc} คือ จำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสได้สำเร็จ ในช่วงเวลาที่กำหนด

T_{all} คือ จำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสในช่วงเวลาที่กำหนด

4) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) คือ ค่าใช้จ่ายในการเรียกใช้บริการของเซอร์วิส ซึ่งในบางบริการของเว็บเซอร์วิสอาจมีการเก็บค่าบริการ วัดจากอัตราการเก็บค่าบริการในการให้บริการเซอร์วิสของผู้ให้บริการในกลุ่มเดียวกัน สูตรการคำนวณแสดงดังสมการที่ 2.4

$$C = 1 - \frac{P}{\sum P_i} \tag{2.4}$$

โดยที่ C คือ ค่าความน่าจะเป็นของค่าใช้จ่ายในการเรียกใช้บริการ

P คือ อัตราการเก็บค่าบริการในการให้บริการเซอร์วิส

$\sum P_i$ คือ ผลรวมของอัตราการเก็บค่าบริการในการให้บริการเซอร์วิสของผู้ให้บริการทั้งหมด

5) ความสำเร็จในการนำส่งข้อมูล (Throughput) คือความสำเร็จในการนำส่งข้อมูลจากจุดหนึ่งไปยังอีกจุดหนึ่งในช่วงเวลาที่กำหนด มีหน่วยเป็นบิตต่อวินาที ซึ่งเป็นจำนวนทรานแซคชัน (Transaction) หรือการร้องขอ (Request) ที่ถูกสร้างขึ้นหรือทำงานได้ในช่วงเวลาการทดสอบหนึ่ง ๆ หน่วยที่นิยมใช้สำหรับวัดค่า Throughput คือ TPS (Transaction Per Second) นั่นคือจำนวนของทราน

แซคชันที่สามารถทำงานได้ใน 1 วินาที ค่าคะแนนคุณภาพของปัจจัยนี้จะบ่งบอกได้ถึงประสิทธิภาพของการให้บริการเว็บเซอร์วิสโดยเฉพาะในเรื่องขนาดของช่องทางที่ใช้ในการสื่อสารข้อมูลหรือแบนด์วิธ เพราะการบริการบางอย่างจำเป็นต้องมีแบนด์วิธเพียงพอจึงจะสามารถใช้งานได้อย่างมีประสิทธิภาพ สูตรการคำนวณค่าแสดงดังสมการที่ 2.5

$$Tp = \frac{N_{suc}}{T_{Total}} \qquad (2.5)$$

$$Tpn = \frac{Tp}{max(Tp)}$$

โดยที่ Tp คือ ค่า Throughput หน่วยเป็น invoke/second

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

T_{total} คือ ค่า ช่วงเวลาการทดสอบ หน่วย second

Tpn คือ ค่า Throughput normalize

max(Tp) คือ ค่า Throughput ที่มีค่าสูงสุด

2.6.3 การคำนวณค่าคุณภาพการให้บริการของเว็บเซอร์วิส

การคำนวณค่าคุณภาพการให้บริการของเว็บเซอร์วิสโดยใช้สมการถ่วงน้ำหนักคำนวณจากการรวมค่าคุณภาพการให้บริการของทั้ง 5 ปัจจัย ได้แก่ เวลาตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ สภาพพร้อมใช้งาน ความน่าเชื่อถือ ค่าใช้จ่ายในการเรียกใช้บริการ และความสำเร็จในการนำส่งข้อมูล สูตรการคำนวณค่าแสดงดังสมการที่ 2.6

$$QoS_{Si} = \sum_{xeV} W_x \; QoS_{Si,x} \qquad (2.6)$$

โดยที่ QoS_{si} คือ ค่าคุณภาพการให้บริการของเว็บเซอร์วิส *Si*

$QoS_{si,x}$ คือ ค่าคุณภาพการให้บริการด้านปัจจัย *x* ของเว็บเซอร์วิส *Si*

W_x คือ ค่าน้ำหนักที่ผู้ใช้บริการกำหนดให้กับคุณภาพด้านปัจจัย *x*

V คือ เซตของปัจจัยที่ใช้กำหนดคุณภาพการให้บริการของเซอร์วิส โดยที่ *V = {Response Time, Availability, Reliability, Cost, Throughput}*

2.7 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) คือ ขั้นตอนหรือกระบวนการทำงานที่ถูกนำมาใช้เพื่อกระทำกับข้อมูลจำนวนมาก ในลักษณะของการสกัดข้อมูล (Data Extraction) เพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นทำให้ได้ความรู้หรือสารสนเทศที่เป็นประโยชน์ (Useful Information) มีความถูกต้อง (Validity) และสามารถนำไปใช้ให้เกิดผลได้ (Actionability) ซึ่งการทำเหมืองข้อมูลถูกนำมาใช้สำหรับค้นหาความรู้ที่มีอยู่ในฐานข้อมูลหรือฐานความรู้ต่าง ๆ มาตั้งแต่ต้นทศวรรษ 1980 (Han, Kamber and Pei, 2011) โดยรูปแบบของความรู้ที่ค้นพบด้วยอัลกอริทึมต่าง ๆ ของการทำเหมืองข้อมูลนั้นจะมีความหมายที่แตกต่างกันไป ขึ้นอยู่กับผู้ใช้งานว่าจะเลือกนำมิติของผลลัพธ์ที่ได้ไปใช้ประโยชน์ให้เหมาะสมกับสิ่งที่ต้องการอย่างไร สำหรับงานวิจัยนี้จะศึกษาแนวคิดและขั้นตอนวิธีของการทำเหมืองข้อมูล เพื่อนำไปใช้ในการพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว โดยเฉพาะในเรื่องของการจำแนกหมวดหมู่ของเว็บเซอร์วิส (Web Services Classification)

2.7.1 ความหมายของการทำเหมืองข้อมูล

กิตติ ภักดีวัฒนะกุล (2550, หน้า 84) กล่าวไว้ว่า เหมืองข้อมูล หมายถึงการวิเคราะห์ข้อมูลเพื่อแยกประเภท จำแนกรูปแบบและความสัมพันธ์ของข้อมูลจากฐานข้อมูลขนาดใหญ่หรือคลังข้อมูล และนำสารสนเทศที่ได้ไปใช้ในการตัดสินใจ ดังนั้นการทำเหมืองข้อมูลจึงมีความเกี่ยวข้องกับการวิเคราะห์ข้อมูลที่มีการใช้เทคนิคบางประการเพื่อค้นหารูปแบบและเชื่อมโยงความสัมพันธ์ของข้อมูลที่ถูกซ่อนอยู่และไม่เคยมีมาก่อนจากคลังข้อมูล จนกลายเป็นการค้นพบองค์ความรู้ใหม่ (Knowledge Discovery)

นิตยา เกิดประสพ (2547, หน้า 9-10) ได้อธิบายความหมายของการทำเหมืองข้อมูลว่าหมายถึง กระบวนการที่นำมาใช้ในการประมวลผลข้อมูลจำนวนมากโดยอัตโนมัติที่เก็บอยู่ในฐานข้อมูล เพื่อค้นหารูปแบบ (Pattern) ที่ซ่อนอยู่ในนั้น เพื่อให้องค์กรนำไปใช้ในการวางแผนการดำเนินงานหรือการตัดสินใจดำเนินการอย่างใดอย่างหนึ่งต่อไป โดยการประมวลผลนี้จะใช้วิธีการทางสถิติ คณิตศาสตร์ การเรียนรู้ของเครื่องและการรู้จำ

ไวท์เทนและแฟรงค์ (Witten and Frank, 2005, p.5) ได้กล่าวว่า การทำเหมืองข้อมูลหมายถึง วิธีการค้นหารูปแบบ (Patterns) ของข้อมูลในฐานข้อมูลที่เป็นไปโดยอัตโนมัติหรือกึ่งอัตโนมัติก็ได้ แต่จะต้องมีข้อกำหนดในกระบวนการค้นหาข้อมูลในฐานข้อมูล (Knowledge Discovery in Databases) อย่างมีรูปแบบ ซึ่งจะทำให้ได้ผลลัพธ์เป็นข้อค้นพบที่มีความหมายในรูปแบบของมิติต่าง ๆ ซึ่งผู้ใช้งานสามารถเลือกนำผลลัพธ์ที่ได้ไปใช้ประโยชน์ตามความต้องการได้

ฮานและแคมเบอร์ (Han and Kamber, 2001, p.5) ได้นิยามความหมายของการทำเหมืองข้อมูลไว้ว่า คือกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น การทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์ รวมทั้งในด้านเศรษฐกิจและสังคม

จากนิยามข้างต้น ผู้วิจัยได้สรุปความหมายของการทำเหมืองข้อมูลว่าหมายถึงกระบวนการในการวิเคราะห์และประมวลผลเพื่อค้นหาสิ่งที่มีประโยชน์จากฐานข้อมูลขนาดใหญ่เพื่อให้ได้มาซึ่งความรู้หรือสารสนเทศหรือรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลนั้น ด้วยเทคนิควิธีต่าง ๆ เช่น การจำแนกข้อมูล (Classification) การจัดกลุ่มข้อมูล (Clustering) และการค้นหากฎความสัมพันธ์ของข้อมูล (Association Rule) เพื่อนำความรู้หรือรูปแบบความสัมพันธ์ที่ได้ไปใช้ประโยชน์ในด้านต่าง ๆ เช่น การจำแนกหรือจัดกลุ่มข้อมูล การแนะนำข้อมูล การวางแผน การพยากรณ์ และการสนับสนุนการตัดสินใจ เป็นต้น

2.7.2 ขั้นตอนการทำเหมืองข้อมูล

การทำเหมืองข้อมูลเป็นกระบวนการที่นำมาใช้ในการประมวลผลข้อมูลจำนวนมากเพื่อค้นหาความรู้หรือรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลนั้น สามารถอธิบายขั้นตอนของการทำเหมืองข้อมูลได้ดังนี้ (Han, Kamber and Pei, 2011)

1) การทำความสะอาดข้อมูล (Data Cleaning) เป็นขั้นตอนการคัดข้อมูลที่ไม่เกี่ยวข้องออกไป ซึ่งอาจจะเป็นข้อมูลที่เสียหรือนำมาใช้ไม่ได้

2) การรวมข้อมูล (Data Integration) เป็นขั้นตอนการรวบรวมข้อมูลที่มีหลายแหล่งให้เป็นข้อมูลชุดเดียวกัน

3) การคัดเลือกข้อมูล (Data Selection) เป็นขั้นตอนการดึงข้อมูลหรือคัดเลือกข้อมูลที่เกี่ยวข้องกับงานที่ต้องการวิเคราะห์ โดยทำการเลือกข้อมูลขึ้นมาจากแหล่งที่บันทึกไว้

4) การแปลงข้อมูล (Data Transformation) เป็นขั้นตอนของการแปลงข้อมูลหรือทำให้ข้อมูลอยู่ในรูปแบบที่เหมาะสมสำหรับการใช้งาน

5) การทำเหมืองข้อมูล (Data Mining) เป็นขั้นตอนการค้นหารูปแบบที่เป็นประโยชน์จากข้อมูลที่มีอยู่ โดยใช้กระบวนการต่าง ๆ ที่จะทำให้สามารถสกัดรูปแบบหรือความรู้จากข้อมูลที่มีอยู่ให้สามารถนำไปใช้งานได้

6) การประเมินตัวแบบ (Pattern Evaluation) เป็นขั้นตอนการประเมินตัวแบบที่ได้จากการทำเหมืองข้อมูลว่าสามารถนำไปใช้เป็นตัวแทนความรู้ได้หรือไม่

7) การนำเสนอความรู้ (Knowledge Representation) เป็นขั้นตอนการนำเสนอความรู้ที่ค้นพบ โดยใช้เทคนิคต่าง ๆ เพื่อนำเสนอเพื่อให้กับผู้ใช้งาน

2.7.3 ลักษณะงานของการทำเหมืองข้อมูล

ผลลัพธ์ที่ได้จากกระบวนการทำเหมืองข้อมูล สามารถนำไปประยุกต์ใช้กับงานลักษณะต่าง ๆ ตามที่ผู้ใช้ต้องการได้อย่างหลากหลาย ตามแนวคิดของแบรี่และรินนอฟ (Berry and Linoff, 2004, p.8) ได้แบ่งลักษณะงานของการทำเหมืองข้อมูลออกเป็น 6 ด้าน ดังนี้

1) งานจำแนกข้อมูล (Classification) เป็นลักษณะงานที่ใช้หลักการจำแนกข้อมูล เพื่อจัดหมวดหมู่และจัดลำดับข้อมูลต่าง ๆ ที่มีการแบ่งกลุ่มอยู่แล้ว ให้สามารถนำผลของการจำแนกข้อมูลที่ได้ไปใช้ประโยชน์ ทำให้เกิดความเข้าใจ ติดต่อสื่อสาร รวมทั้งเข้าถึงได้ง่าย ซึ่งเป็นเทคนิควิธีดั้งเดิมและมีการใช้งานกว้างขวางที่สุด

2) งานจัดกลุ่ม (Clustering) เป็นลักษณะงานที่ใช้วิธีการจัดกลุ่มข้อมูลที่มีความเกี่ยวข้องสัมพันธ์กัน หรือมีลักษณะที่เหมือน ๆ กันให้อยู่ในกลุ่มเดียวกัน เป็นกระบวนการ ที่ไม่ต้องอาศัยการจำแนกข้อมูลไว้ล่วงหน้า มีหลักการทำงานที่แตกต่างกับงานจำแนกข้อมูลอย่างเด่นชัด โดยงานจัดกลุ่มข้อมูลจะต้องทำการค้นหาว่าข้อมูลจะถูกนำมารวมกลุ่มหรือรวมตัวกันบนพื้นฐานของความคล้ายกันของข้อมูลและจัดเป็นกลุ่มต่าง ๆ ได้อย่างไร ซึ่งจะต้องแบ่งกลุ่มหรือจัดกลุ่มให้เสร็จสมบูรณ์ก่อน จึงจะสามารถนาไปใช้ประโยชน์ได้ (Witten and Frank, 2005, p.136)

3) งานอธิบายและบอกรายละเอียดข้อมูล (Description and Profiling) เป็นงานที่มุ่งเน้นการนำผลที่ได้จากการทำเหมืองข้อมูลไปใช้ประโยชน์ในรูปแบบของการอธิบายหรือบอกรายละเอียดสิ่งที่เกิดขึ้นในฐานข้อมูลที่มีความสลับซับซ้อน เพื่อให้เกิดความเข้าใจเกี่ยวกับข้อมูลที่มีอยู่มากขึ้น เช่น การอธิบายและวิเคราะห์ความสัมพันธ์ระหว่างข้อมูลประวัติลูกค้ากับข้อมูลผลิตภัณฑ์ที่เกิดขึ้นร่วมกัน ด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) หรือกฎความสัมพันธ์ (Association Rule) ก็จะทำให้สามารถนำไปใช้ในประโยชน์ทางการค้าได้ เป็นต้น

4) งานประเมินค่า (Estimation) เป็นงานที่เกี่ยวข้องกับการนำผลที่ได้จากการทำเหมืองข้อมูลไปประยุกต์ใช้เพื่อประเมินค่าสิ่งต่าง ๆ ซึ่งจะทำให้ได้ผลลัพธ์ที่มีประโยชน์และอาจจะเป็นสิ่งที่ไม่ได้คาดคิดมาก่อน โดยใช้วิธีการสร้างแบบจำลองหรือตัวแบบจากเทคนิควิธีการต่าง ๆ จากนั้นนำแบบจำลองที่ได้ไปทำการป้อนข้อมูลบางอย่างเข้าไป ซึ่งก็จะทำให้สามารถประเมินมูลค่าของสิ่งที่มีความสนใจหรือต้องการได้

5) งานค้นหากฎความสัมพันธ์ (Association Rule Discovery) เป็นลักษณะงานที่เกี่ยวข้องกับการค้นหารูปแบบความสัมพันธ์ ที่อยู่ในรูปแบบของชุดข้อมูล (Item Sets) กับเหตุการณ์ที่ตามมา (Subsequences) และโครงสร้างของส่วนย่อย (Substructures) ซึ่งจะได้ผลลัพธ์เป็นลักษณะ

ความสัมพันธ์ของชุดข้อมูลที่พบหรือปรากฏร่วมกันบ่อย ๆ (Han, Kamber and Pei, 2011, pp.228-229) ซึ่งเป็นงานของการทำเหมืองข้อมูลที่ได้รับความนิยมนำไปใช้ในวงการธุรกิจ โดยเฉพาะในส่วนของการวิเคราะห์พฤติกรรมในการซื้อสินค้าของลูกค้า

6) งานคาดการณ์ (Prediction) เป็นลักษณะงานที่เกี่ยวข้องกับกระบวนการสร้างรูปแบบเพื่อคาดการณ์หาค่าที่ต้องการจากข้อมูลที่มีอยู่แล้ว ซึ่งเป็นการคาดการณ์โดยอาศัยการหาค่าที่ต้องการจากข้อมูลที่มีอยู่ในอดีต จึงทำให้ได้ค่าที่ต้องการออกมาเป็นตัวเลข โดยผลลัพธ์ที่ได้จากวิธีการนี้จะทำให้สามารถคาดการณ์เหตุการณ์ต่าง ๆ ที่จะเกิดขึ้นในอนาคตได้ว่าจะมีแนวโน้มเป็นอย่างไร ตัวอย่างของการไปประยุกต์ใช้งาน เช่น การคาดการณ์แนวโน้มของตลาดหุ้น สภาพอากาศ วิเคราะห์ความหนาแน่นของช่องสัญญาณในระบบสื่อสารขนาดใหญ่ เป็นต้น

2.7.4 การจำแนกหมวดหมู่ของเว็บเซอร์วิสโดยใช้วิธีการทำเหมืองข้อมูล

งานวิจัยนี้ใช้เทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสอยู่หลายเทคนิคเพื่อนำมาเปรียบเทียบว่าเทคนิคใดจะเหมาะสมกับข้อมูลเว็บเซอร์วิสที่ให้บริการที่เกี่ยวข้องด้านการท่องเที่ยวมากที่สุด และมีการประเมินผลว่าเทคนิคใดให้ความถูกต้องในการจำแนกหมวดหมู่ของเว็บเซอร์วิสได้ถูกต้องมากที่สุด ซึ่งเทคนิคต่าง ๆ มีดังนี้

1) เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT)

เทคนิควิธีต้นไม้ตัดสินใจเป็นวิธีการทำเหมืองข้อมูลวิธีหนึ่งที่ได้รับความนิยมในการจำแนกหมวดหมู่ของเอกสาร โดยมีกลุ่มข้อมูลหรือเป้าหมายที่ต้องการอย่างชัดเจน เป็นกระบวนการที่นำมาใช้ในการค้นหาตัวแบบหรือแพทเทิร์น (Pattern) ของข้อมูลเพื่อให้สามารถจำแนกข้อมูลออกเป็นกลุ่มหรือที่เรียกว่าคลาส (Class) ซึ่งอยู่ในรูปแบบของลักษณะโครงสร้างต้นไม้ภายในประกอบไปด้วยโหนด (Node) และกิ่ง (Branch) แต่ละโหนดจะถูกแทนด้วยคุณลักษณะ (Feature) กระบวนการจำแนกข้อมูลสามารถแบ่งออกเป็น 2 ส่วนใหญ่ ๆ คือ ส่วนแรกเป็นการสร้างแบบจำลองหรือตัวแบบโดยนำข้อมูลชุดฝึกหัด (Training Data) ที่จัดเตรียมไว้มาผ่านกระบวนการค้นหาตัวแบบด้วยอัลกอริทึมต่าง ๆ โดยตัวแบบที่ได้จะเรียกว่า ตัวจำแนก (Classifier) ซึ่งจะถูกนำไปใช้สำหรับจำแนกข้อมูลต่อไป สำหรับส่วนที่สอง เป็นการนำตัวจำแนกที่ได้มาจากส่วนแรกไปทดสอบเพื่อวัดค่าความแม่นตรงของตัวแบบกับข้อมูลชุดทดสอบ (Testing Data) รวมทั้งนำไปใช้ประโยชน์ในการพยากรณ์หรือทำนายข้อมูลใหม่ โดยเกณฑ์การเลือกคุณลักษณะเพื่อนำมาเป็นโหนดของต้นไม้นั้นมาจากการคำนวณค่าเกนสารสนเทศ (Information Gain) โดยพิจารณาคุณลักษณะที่มีค่าเกนสารสนเทศหรือมีค่าเอ็นโทรพี (Entropy) ต่ำแสดงว่าคุณลักษณะนั้นมีประสิทธิภาพในการจำแนกหมวดหมู่สูง เมื่อทำการประมวลผลข้อมูลโดยวิธีนี้พบว่ากฎการจำแนกที่ได้จากเทคนิควิธีต้นไม้ตัดสินใจนั้น สามารถเข้าใจได้ง่ายและไม่ยุ่งยากในการนำไปใช้งาน รวมทั้งมีกลไกของการ

เรียนรู้ในกระบวนการสร้างตัวแบบที่รวดเร็วกว่ากลไกของเทคนิควิธีอื่น ๆ จึงส่งผลให้เทคนิคต้นไม้ตัดสินใจถูกนำไปประยุกต์ใช้งานอย่างหลากหลาย (Quinlan, 1993) เทคนิคต้นไม้ตัดสินใจสามารถแสดงได้ดังสมการที่ 2.7

โดยที่ T แทน เซตของข้อมูลชุดฝึกหัด (Training Data)

X แทน แอททริบิวท์ที่ถูกเลือกให้เป็นตัวตรวจสอบเพื่อจัดกลุ่มข้อมูล

สามารถหาค่า Gain ได้ดังสมการที่ 2.7 ดังนี้ (Han, Kamber and Pei, 2011)

$$Gain(X) = Info(T) - Info_x(T)\ _{Bits} \tag{2.7}$$

โดยที่ *Info(T)* คือ ฟังก์ชันที่ระบุปริมาณข้อมูลที่ต้องการเพื่อให้สามารถจำแนกคลาสของข้อมูลได้

$$Info(T) = -\sum_{i=1}^{k} \left[\frac{freq(C_i,T)}{|T|}\right] * \ log_2 \left[\frac{freq(C_i,T)}{|T|}\right] Bits \tag{2.8}$$

โดยที่ $|T|$ คือ จำนวนข้อมูลทั้งหมดในเซตของข้อมูลชุดฝึกหัด

$freq(C_i, T)$ คือ ความถี่ที่ข้อมูลใน T ปรากฏในคลาส C_i

$Info_x(T)$ คือ ฟังก์ชันที่ระบุปริมาณข้อมูลที่ต้องการ โดยใช้แอททริบิวท์ X เป็นตัวตรวจสอบเพื่อจำแนกคลาสกลุ่มของข้อมูลได้

$$Info_x(T) = \sum_{i=1}^{n} \left[\frac{|T_i|}{|T|}\right] * \ info(T_i)\ Bits \tag{2.9}$$

โดยที่ i คือ จำนวนค่าที่เป็นไปได้ของแอททริบิวท์

$X|T_i|$ คือ จำนวนข้อมูลที่มีค่า $X = i$

2) เทคนิควิธีเคเนียร์เรสท์เนเบอร์ (*k*-Nearest Neighbor : *k*-NN)

เทคนิควิธีเคเนียร์เรสท์เนเบอร์หรือเทคนิควิธีการเพื่อนบ้านใกล้ที่สุด เป็นวิธีที่ใช้ในการจัดแบ่งหมวดหมู่ข้อมูล โดยเทคนิคนี้จะตัดสินใจว่า คลาสใดที่จะแทนเงื่อนไขหรือกรณีใหม่ ๆ ได้บ้าง โดยการตรวจสอบจำนวนบางจำนวนของกรณีหรือเงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด โดยจะหาผลรวมของจำนวนเงื่อนไข หรือกรณีต่าง ๆ สำหรับแต่ละคลาส และกำหนดเงื่อนไข

ใหม่ ๆ ให้คลาสที่เหมือนกันกับคลาสที่ใกล้เคียงกันมากที่สุด เทคนิควิธีนี้เป็นเทคนิคที่แตกต่างจากเทคนิคอื่นคือไม่ได้ใช้ข้อมูลฝึกหัด (Training Data) ในการสร้างแบบจำลอง แต่จะใช้ข้อมูลนั้นมาเป็นตัวแบบจำลองเลย ในการใช้งานเทคนิควิธีนี้ต้องระบบค่าตัวเลขจำนวนเต็มบวกให้กับ *k* ด้วย ซึ่งค่านี้จะเป็นตัวบอกจำนวนของกรณีที่จะต้องค้นหาในการทำนายกรณีใหม่ (Shakhnarovich, Darrell and Indyk, 2006)

การนำเทคนิคของเคเนียร์เรสท์เนเบอร์ไปใช้นั้น เป็นการหาระยะห่างระหว่างแต่ละแอททริบิวท์ (Attribute) ในข้อมูล จากนั้นคำนวณค่าออกมา ซึ่งวิธีนี้จะเหมาะสำหรับข้อมูลแบบตัวเลข แต่ตัวแปรที่เป็นค่าแบบไม่ต่อเนื่องนั้นก็สามารถทำได้ เพียงแต่ต้องการการจัดการแบบพิเศษเพิ่มขึ้น จากนั้นต้องมีวิธีในการรวมค่าระยะห่างของแอททริบิวท์ (Attribute) ทุกค่าที่วัดมาได้ เมื่อสามารถคำนวณระยะห่างระหว่างเงื่อนไขหรือกรณีต่าง ๆ ได้ จากนั้นเลือกชุดของเงื่อนไขที่ใช้จัดคลาส มาเป็นฐานสำหรับการจัดคลาสในเงื่อนไขใหม่ ๆ ได้แล้วจะตัดสินได้ว่าขอบเขตของจุดข้างเคียงที่ควรเป็นนั้น ควรมีขนาดใหญ่เท่าไร และอาจมีการตัดสินใจได้ด้วยว่าจะนับจำนวนจุดข้างเคียงตัวมันได้อย่างไร ประสิทธิภาพของขั้นตอนเทคนิควิธีเคเนียร์เรสท์เนเบอร์นี้ จะขึ้นอยู่กับจำนวนระยะห่าง การอธิบายระหว่างข้อมูลทั้งคู่ ที่สามารถแบ่งแยกอย่างมีประสิทธิภาพระหว่างข้อมูลปกติ และข้อมูลผิดปกติ การอธิบายจำนวนระยะห่างระหว่างขอมูลเป็นความท้าทายอย่างมากเมื่อข้อมูลมีความซับซ้อน เช่น ข้อมูลกราฟ และข้อมูลแบบลำดับ เป็นต้น (Shakhnarovich, Darrell and Indyk, 2006) การคำนวณหาระยะห่างระหว่างข้อมูลแต่ละตัวจะใช้วิธีการวัดระยะทางแบบยูคลิเดียน (Euclidian Distance)

3) เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB)

เป็นการทำเหมืองข้อมูลวิธีหนึ่งที่เหมาะในการนำไปใช้งานทางด้านการจำแนกหมวดหมู่ของข้อมูล หลักการของเทคนิควิธีนี้เป็นวิธีการเรียนรู้ที่ใช้หลักการของความน่าจะเป็น (Conditional Proabability) ซึ่งมีพื้นฐานมาจากทฤษฎีของเบย์ (Bayes theorem) เข้ามาช่วยในการเรียนรู้ จุดมุ่งหมายเพื่อต้องการสร้างโมเดลที่อยู่ในรูปแบบความน่าจะเป็น ซึ่งเป็นค่าที่บันทึกได้จากการสังเกต จากนั้นนำโมเดลมาหาว่าสมมติฐานใดถูกต้องที่สุดโดยใช้ความน่าจะเป็นเข้ามาช่วย การจำแนกหมวดหมู่ข้อมูลวิธีนี้เป็นวิธีที่มีประสิทธิภาพวิธีหนึ่งเหมาะกับกรณีของเซตตัวอย่างมีจำนวนมากและคุณสมบัติของตัวอย่างไม่ขึ้นต่อกัน มีการนำไปประยุกต์ใช้งานในด้านการจำแนกประเภทข้อความ (Text Classification) และมีอัลกอริทึมในการทำงานที่ไม่ซับซ้อนเหมือนวิธีการอื่น ๆ (Domingos and Pazzani, 1997) โดยที่ *A* คือ แอตทริบิวต์ (Attribute) และ *C* คือ ค่าคลาส (Class) ดังสมการที่ 2.10

$$P(C|A) = \frac{P(A|C)*P(C)}{P(A)} \quad (2.10)$$

โดยที่ P($C|A$) คือ ค่าความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์เป็น A จะมีคลาส C

P($A|C$) คือ ค่าความน่าจะเป็นที่ข้อมูลชุดฝึกหัด (Training Data) ที่มีคลาส C และมีแอตทริบิวต์ A โดยที่ $A = a1 \cap a2 \ldots \cap aM$ โดยที่ M คือ จำนวนแอตทริบิวต์ในชุดข้อมูลฝึกหัด (Training Data)

P(C) คือ ค่าความน่าจะเป็นของคลาส C

4) เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM)

เป็นเทคนิควิธีที่สามารถนำมาช่วยแก้ปัญหาในการจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด แนวความคิดของเทคนิควิธีนี้เกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (Feature Space) จากนั้นจึงหาเส้นที่ใช้แบ่งข้อมูลทั้งสองออกจากกันโดยจะสร้างเส้นแบ่ง (Hyperplane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันนั้น เส้นตรงใดเป็นเส้นที่ดีที่สุด สำหรับพื้นฐานเดิมของเทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีนนี้ ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยการนำฟังก์ชันเคอร์เนล (Kernal Function) มาใช้ (Cristianini and Shawe-Taylor, 2000)

การจำแนกข้อมูลบนระนาบหลายมิติ จะใช้ส่วนการเลือกที่มีความเหมาะสมที่สุด เรียกว่า โครงสร้างในการคัดเลือก (Feature Selection) ซึ่งโครงสร้างในการคัดเลือกมาจากข้อมูลที่สอนให้ระบบเรียนรู้ จำนวนเซตของโครงสร้างที่ใช้อธิบายในกรณีหนึ่ง เรียกว่า เวกเตอร์ (Vector) ดังนั้นจุดมุ่งหมายของตัวแบบซัพพอร์ทเวกเตอร์แมชชีนคือ แบ่งแยกกลุ่มของเวกเตอร์ในกรณีนี้ด้วยหนึ่งกลุ่มของตัวแปรเป้าหมายที่อยู่ข้างหนึ่งของระนาบ และกรณีของกลุ่มอื่นที่อยู่ทางระนาบต่างกัน ซึ่งเวกเตอร์ที่อยู่ข้างระนาบหลายมิติทั้งหมดเรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vectors)

5) เทคนิควิธีเคมีนส์ (K-means)

เป็นเทคนิคหนึ่งที่ใช้ในการจัดกกลุ่มข้อมูล จัดอยู่ในเทคนิคแบบการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) โดยการจัดกลุ่มจะแบ่งข้อมูลออกเป็นกลุ่มหรือคลัสเตอร์ (Cluster) โดยจะจัดให้ข้อมูลที่มีความคล้ายคลึงกันอยู่ในคลัสเตอร์เดียวกัน และข้อมูลที่อยู่ต่างคลัสเตอร์กันจะมีความคล้ายคลึงกันน้อยที่สุด ในการจัดกลุ่มข้อมูลมีการใช้ค่าเฉลี่ยของข้อมูลที่ถูกจัดให้อยู่ในคลัส

เตอร์เดียวกันเป็นตัวแทนของทุกข้อมูลในคลัสเตอร์นั้น ขั้นตอนการจัดกลุ่มเริ่มต้นจากการรับค่าพารามิเตอร์ k ซึ่งค่านี้คือจำนวนคลัสเตอร์ที่ต้องการค้นหา จากนั้นขั้นตอนวิธีจะทำการสุ่มเลือกข้อมูลเริ่มต้นจำนวน k ชุด ซึ่งแต่ละชุดที่ได้มานั้นจะเป็นจุดศูนย์กลางเริ่มต้น (Centroid) ของแต่ละคลัสเตอร์ จากนั้นทำการจัดกลุ่มให้กับข้อมูลที่เหลือ ข้อมูลจะถูกจัดให้อยู่ในคลัสเตอร์เดียวกันเมื่อข้อมูลนั้นมีความคล้ายกับตัวแทนของคลัสเตอร์นั้นมากที่สุด จากนั้นจึงทำการคำนวณหาค่าเฉลี่ยของคลัสเตอร์ใหม่ และดำเนินกระบวนการเดียวกันกับข้อมูลที่เหลือต่อไป จนกระทั่งทุกข้อมูล ถูกจัดกลุ่มอย่างสมบูรณ์และข้อมูลไม่มีการเปลี่ยนกลุ่มอีกต่อไป การทำงานของเทคนิควิธีเคมีนส์จะมีประสิทธิภาพสูงก็ต่อเมื่อข้อมูลเกาะกลุ่มกันหนาแน่น แต่ละกลุ่มแยกจากกันอย่างชัดเจน และความหนาแน่นของข้อมูลในแต่ละกลุ่มใกล้เคียงกัน จุดเด่นของเทคนิควิธีเคมีนส์ คือง่ายและสามารถใช้ได้กับข้อมูลหลายประเภท และยังมีประสิทธิภาพในด้านความเร็ว แต่จุดด้อยของเทคนิควิธีเคมีนส์ก็พบว่ายังไม่เหมาะสมกับข้อมูลทุกประเภท และไม่สามารถจัดการกลุ่มที่มีรูปร่างไม่เป็นรูปทรงกลมหรือกลุ่มที่มีขนาดหรือความหนาแน่นแตกต่างกันได้ นอกจากนี้เทคนิควิธีเคมีนส์ยังถูกจำกัดสำหรับข้อมูลที่มีตัวแทนข้อมูลที่คลุมเครือหรือไม่ชัดเจน (Kanungo et al., 2002)

2.8 เทคนิคการวัดความคล้ายคลึงกันแบบโคไซน์ (Cosine Similarity)

ในการวัดความคล้ายคลึงของเว็บเซอร์วิสที่อยู่ในกลุ่มบริการเดียวกันใช้วิธีการหาค่าความคล้ายคลึงตามทฤษฎีแบบจำลองเวกเตอร์สเปซ (Vector Space Model: VSM) คือการวัดความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) เป็นการวัดความคล้ายคลึงเชิงมุม ซึ่งใช้การแทนเอกสารด้วยระบบเวกเตอร์และหลักการทางสถิติในการวัด (Salton, 1988; Sidorov et al., 2014) แสดงดังสมการที่ 2.11

$$\text{similarity} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{2.11}$$

โดยที่ similarity คือ ค่าคะแนนความคล้ายคลึงกันของเว็บเซอร์วิส

A_i คือ เวกเตอร์ของเว็บเซอร์วิสหลัก ค่าที่เป็นไปได้คือ 0 หรือ 1

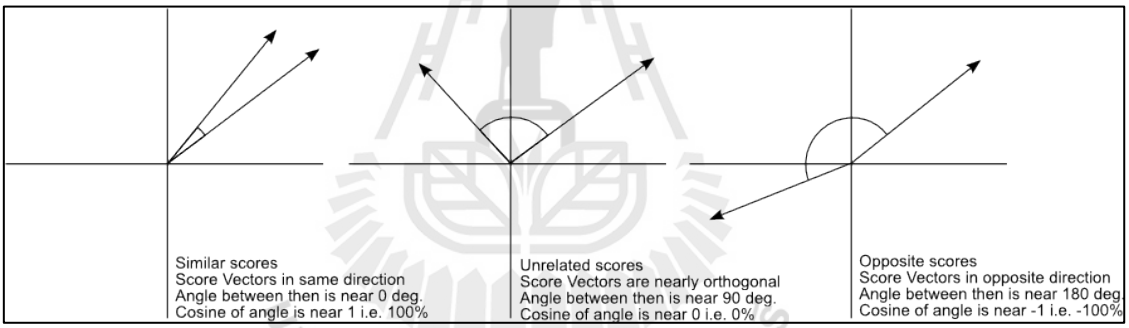
B_i คือ เวกเตอร์ของเว็บเซอร์วิสรอง ค่าที่เป็นไปได้คือ 0 หรือ 1

$\sum_{i=1}^{n} A_i B_i$ คือ ค่าผลรวมของเวกเตอร์ของเว็บเซอร์วิสหลักและรอง

$\sqrt{\sum_{i=1}^{n} A_i^2}$ คือ ค่ารากที่สองของผลรวมเวกเตอร์ของเว็บเซอร์วิสหลักยกกำลัง 2

$\sqrt{\sum_{i=1}^{n} B_i^2}$ คือ ค่ารากที่สองของผลรวมเวกเตอร์ของเว็บเซอร์วิสรองยกกำลัง 2

เทคนิคการวัดความคล้ายคลึงกันแบบโคไซน์ (Cosine similarity) เป็นวิธีการหาคล้ายคลึงกันจากค่าความต่างของมุมของวัตถุสองวัตถุที่เกิดขึ้นบนพื้นที่เวกเตอร์ วิธีการนี้เป็นที่นิยมและมีประสิทธิภาพสูงในการวัดความคล้ายคลึงระหว่างวัตถุสองวัตถุ หรือระหว่างเอกสารทั้งสองเอกสาร และถูกนำมาประยุกต์ใช้กับศาสตร์ทางด้านการค้นคืนข้อมูลอย่างแพร่หลาย เนื่องจากวิธีการนี้จะมีประสิทธิภาพในกรณีที่เอกสารสองเอกสารมีความยาวไม่เท่ากันหรือทำให้มีความยุติธรรมต่อเอกสารที่สั้นกว่า ซึ่งค่าความคล้ายคลึงกันแบบโคไซน์นี้จะมีค่าอยู่ระหว่าง 0-1 เท่านั้น โดยแต่ละเอกสารจะถูกแทนด้วยเวกเตอร์ซึ่งเก็บค่าน้ำหนักคำแต่ละคำในเอกสารนั้น การเปรียบเทียบความคล้ายคลึงของเอกสารจะพิจารณาจากมุมโคไซน์ของมุมระหว่างสองเวกเตอร์ของเอกสาร ทั้งนี้ค่าความคล้ายคลึงสูงสุดที่วัดด้วยวิธีนี้จะมีค่าเท่ากับ 1 ความหมายคือเวกเตอร์ทั้งสองทำมุมระหว่างกัน 0 องศา นั้นคือเวกเตอร์ทั้งสองมีทิศทางเดียวกันหรือมีความคล้ายคลึงกัน 100 เปอร์เซ็นต์ และถ้าค่าความคล้ายคลึงมีค่าเท่ากับ 0 เวกเตอร์ทั้งสองจะทำมุมระหว่างกัน 90 องศา หมายถึงเวกเตอร์ทั้งสองหรือเอกสารทั้งสองไม่มีความคล้ายคลึงกันเลย แสดงดังรูปที่ 2.12



รูปที่ 2.12 ค่าความคล้ายคลึงแบบโคไซน์ที่แสดงถึงความแตกต่างระหว่างสองเอกสาร (Perone, www, 2013)

ในงานวิจัยนี้ผู้วิจัยได้นำเทคนิคการวัดความคล้ายคลึงแบบโคไซน์มาใช้ในการเปรียบเทียบหาความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่มบริการ โดยพิจารณาจากชื่อของเว็บเซอร์วิส (Web Service Name) คำอธิบาย (Description) ชื่อโอเปอเรชัน (Operation Name) ข้อมูลเข้า (Input) และข้อมูลออก (Output) ซึ่งองค์ประกอบของเว็บเซอร์วิสทั้งหมดนี้เปรียบเสมือนเวกเตอร์ที่จะเข้าสู่กระบวนการคำนวณความคล้ายคลึงกันตามเทคนิคการวัดความคล้ายคลึงแบบโคไซน์

2.9 การจัดลำดับข้อมูล (Ranking)

การจัดลำดับ (Ranking) เป็นการนำข้อมูลที่มีลักษณะแตกต่างกันมาจัดเรียงกัน ซึ่งการเรียงลำดับจะเป็นที่น่าเชื่อถือได้เมื่อการเรียงลำดับนั้น ๆ ใช้การพิจารณาคุณลักษณะเดียวที่มีการนิยามไว้อย่างชัดเจน ซึ่งถ้าทำการเรียงอันดับโดยพิจารณาหลายคุณลักษณะไปพร้อม ๆ กัน นอกจากจะทำได้ยากแล้วยังอาจจะมีผลต่อความเชื่อมั่นด้วย ดังนั้นในการจัดลำดับของข้อมูลจึงจำเป็นต้องมีการสร้างหลักเกณฑ์ต่าง ๆ เพื่อประกอบการตัดสินใจ (บุญชม ศรีสะอาด, 2543, หน้า 56)

วิฑูรย์ ตันศิริคงคล (2542, หน้า 179-185) ได้กล่าวถึงการจัดลำดับว่าเป็นวิธีการวัดเชิงสุทธิ ซึ่งหมายความว่า เป็นการจัดลำดับทางเลือกต่าง ๆ ที่เป็นอิสระต่อกันทีละตัว โดยอาศัยลำดับความเข้มข้นสำหรับแต่ละเกณฑ์ของการตัดสินใจ ซึ่งหลักการของกระบวนการตัดสินใจในส่วนของการจัดกลุ่มองค์ประกอบของปัญหา ได้แก่ การสร้างเกณฑ์หลักในการตัดสินใจ และเกณฑ์รองในการตัดสินใจ รวมถึงการเชื่อมโยงอิทธิพลระหว่างองค์ประกอบ โดยมีขั้นตอนในการจัดลำดับประกอบด้วย 1) ให้ลำดับความสำคัญของเกณฑ์หลักและเกณฑ์รอง 2) กำหนดมาตรฐานในรูปของความเข้มข้น เช่น ดีมาก ดี ปานกลาง น้อย ต่ำกว่ามาตรฐาน เป็นต้น 3) กำหนดลักษณะที่เป็นรูปธรรมให้กับมาตรฐานโดยการหาค่าความเข้มข้นของเกณฑ์ และ 4) ให้คะแนนตามเกณฑ์ในแต่ละทางเลือก

มาล์ซิวสกิ (Malczewski, 1999, pp.178-179) ได้อธิบายเกี่ยวกับหลักเกณฑ์การหาค่าน้ำหนักสำหรับจัดลำดับข้อมูลด้วยเทคนิควิธีหาผลรวม (Rank Sum) ว่าเป็นเทคนิคการจัดลำดับข้อมูลที่ใช้หลักการปรับค่าน้ำหนักของข้อมูลให้เป็นมาตรฐานเดียวกัน เริ่มจากการหาค่าส่วนกลับของลำดับความสำคัญของข้อมูลแต่ละรายการด้วยสูตร n-rj+1 ซึ่งจะทำให้สามารถคำนวณหาค่าผลรวมของค่าส่วนกลับของลำดับความสำคัญของข้อมูลทั้งหมดได้ จากนั้นจึงนำค่าผลรวมที่ได้ไปทำการปรับเป็นค่าน้ำหนักของข้อมูลแต่ละอันดับให้เป็นมาตรฐานเดียวกัน โดยมีสูตรการคำนวณแสดงดังสมการที่ 2.12

$$RankSum(W_i) = \frac{n-r_j+1}{\sum(n-r_k+1)} \tag{2.12}$$

โดยที่ *Wi* คือ ค่าน้ำหนักของข้อมูล
rj คือ ลำดับความสำคัญของข้อมูล
rk คือ ผลรวมของค่าของข้อมูล
n คือ จำนวนข้อมูลทั้งหมด

จากการที่ผู้วิจัยได้ทบทวนวรรณกรรมที่เกี่ยวข้องกับการจัดลำดับข้อมูล พบว่าการทำดัชนีแบบกำหนดค่าน้ำหนักที่มีการจัดทำอย่างมีคุณภาพตามเกณฑ์มาตรฐาน นอกจากจะทำให้ผลการ

สืบค้นที่ตรงกับความต้องการมากขึ้นเพราะค่าน้ำหนักของดัชนีจะทำให้ได้เลือกแสดงผลเฉพาะข้อมูลที่มีค่าน้ำหนักดัชนีสูงกว่าหรือเท่ากับค่าน้ำหนักขั้นต่ำที่กำหนดแล้ว การกำหนดค่าน้ำหนักยังสามารถแก้ปัญหาจำนวนผลลัพธ์จากการสืบค้นที่มีมากเกินไปได้ เพราะการทำดัชนีแบบกำหนดค่าน้ำหนักสามารถช่วยเรียงลำดับข้อมูลที่มีค่าคะแนนจากมากที่สุดไปน้อยที่สุดได้ (สุธรรม อุมาแสงทองกุล, 2541, หน้า 12) ซึ่งจะเห็นได้ว่า เทคนิควิธีการจัดลำดับโดยการหาผลรวม (Rank Sum) นอกจากจะเป็นเทคนิคที่ง่าย นิยมใช้อย่างแพร่หลาย และเหมาะสำหรับจัดอันดับข้อมูลแล้ว (Malczewski, 1999, p.190) ยังพบว่าเทคนิคการจัดลำดับโดยการหาผลรวมนี้มีคุณสมบัติที่สามารถนำไปประยุกต์ใช้สำหรับการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวได้ เพราะสามารถนำเทคนิคนี้มาใช้ในการจัดอันดับของข้อมูลได้ เช่น การสร้างหลักเกณฑ์เพื่อกำหนดค่าน้ำหนักให้กับทุกตำแหน่งของคำศัพท์ตามโครงสร้างคลังคำแบบสคอส (SKOS) กำหนดค่าน้ำหนักกับตำแหน่งของคำค้นจากผู้ใช้ และกำหนดค่าน้ำหนักของคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส เป็นต้น

2.10 งานวิจัยที่เกี่ยวข้อง

งานวิจัยเรื่องระบบสืบค้นข้อมูลแบบบูรณาการเชิงความหมายสำหรับข้อมูลภายในองค์กร (มารุต บูรณรัช และคณะ, 2553) จากปัญหาการสืบค้นข้อมูลในองค์กรที่มีปริมาณมากและถูกจัดเก็บหลายรูปแบบ เช่น เอกสาร อีเมล เว็บเพจ ฐานข้อมูล เป็นต้น โดยใช้คำสำคัญหรือคำค้นในการค้นหาข้อมูล ทำให้ผู้ใช้เสียเวลาในการค้นหาและพบเพียงเอกสารที่มีความเกี่ยวข้องแต่ไม่ใช่ข้อมูลที่ต้องการจริง เนื่องจากผลลัพธ์ที่ได้ขาดการบูรณาการและเชื่อมโยงกับข้อมูลที่เกี่ยวข้อง งานวิจัยนี้ใช้เทคนิคการสืบค้นข้อมูลที่ใช้การจัดระเบียบข้อมูลเชิงความหมาย (Semantic-Based Information Organization) สนับสนุนการบูรณาการและจัดระเบียบข้อมูล เพื่อให้ได้ผลลัพธ์จากการสืบค้นที่มีการสรุปสาระสำคัญ และมีการเชื่อมโยงข้อมูลเชิงความหมายหลายรูปแบบประกอบกัน โดยครอบคลุมข้อมูลทั้งชนิดมีโครงสร้าง (Structured Data) และข้อมูลชนิดไร้โครงสร้าง (Unstructured Data) และไม่จำกัดว่าข้อมูลมาจากฐานข้อมูลเดียวหรือหลายฐานข้อมูล โดยใช้การจัดระเบียบที่อิงกับข้อมูลเชิงความหมายทั้งที่มาจากการสกัดสาระสำคัญออกจากฐานข้อมูลโดยอัตโนมัติ โดยข้อมูลที่ผู้ใช้ต้องการมักเป็นสาระสำคัญ เช่น เวลา สถานที่ นิยาม เหตุผล บุคคล ผู้เชี่ยวชาญ สิ่งที่เกี่ยวข้อง เป็นต้น และข้อมูลเชิงความหมายที่สร้างขึ้นใหม่เองในรูปแบบของออนโทโลยีช่วยให้ผู้ใช้สามารถระบุความต้องการข้อมูลได้ชัดเจนมากยิ่งขึ้น โดยใช้ข้อมูลในรูปแบบของ RDF และ OWL และสามารถประยุกต์ใช้ระบบอนุมานอัตโนมัติได้ (Inference Engine) และการสืบค้นจะอิงกับการจัดกลุ่มและความสัมพันธ์ของสิ่งต่าง ๆ ในเชิงความหมายช่วยให้การสืบค้นข้อมูลเชิงความหมายมีประสิทธิภาพมากขึ้น

งานวิจัยเรื่องการกำหนดดัชนีคำสำคัญโดยอัตโนมัติ (Hulth, 2003a; 2003b) เป็นงานวิจัยที่ศึกษาถึงการกำหนดและคัดเลือกคำสำคัญ (Keyword) ให้กับเอกสาร โดยเปรียบเทียบกับการคัดเลือกคำสำคัญโดยอาศัยมนุษย์เป็นผู้กำหนด ซึ่งผู้วิจัยเชื่อว่าการจัดทำฐานดัชนีคำสำคัญ (Keyword Indexing) เป็นส่วนสำคัญที่จะช่วยทำให้การค้นคืนข้อมูลทำงานได้มีประสิทธิภาพมากขึ้น โดยทดลองกับการจัดเก็บเอกสารภายในห้องสมุดของรัฐสภาสวีเดน ซึ่งมีข้อบกพร่องสำหรับการค้นหาคือผู้ใช้จำเป็นจะต้องรู้ว่าหนังสือที่ต้องการค้นหานั้นน่าจะถูกจัดเก็บไว้ในหมวดหมู่ไหน และบ่อยครั้งที่เอกสารเดียวกันถูกเก็บไว้ในหลายหมวดหมู่ จึงได้แก้ไขโดยการจัดทำดัชนีให้กับเอกสาร ซึ่งใช้กำหนดคำสำคัญที่มีจะบ่งชี้หัวข้อและเนื้อหาสำคัญในเอกสารและพบว่าการทำดัชนีโดยอาศัยมนุษย์อาจมีความผิดพลาดได้ เนื่องจากขาดความแม่นยำและเกณฑ์ที่ใช้ในการจัดทำดัชนีของผู้ประเมินแตกต่างกัน ผู้วิจัยจึงได้นำเอาความรู้ในการกำหนดขอบเขตข้อมูล (Domain Knowledge) มาใช้ในการตัดคำสำคัญโดยอัตโนมัติเพื่อจัดทำดัชนีให้กับเอกสารที่ถูกจัดเก็บในห้องสมุด มีการทดลองโดยรวบรวมเอกสารจากห้องสมุดที่ได้มีการจัดทำดัชนีโดยอาศัยผู้เชี่ยวชาญ และใช้หลักการการเรียนรู้จดจำของเครื่อง (Machine Learning Algorithm) ในการพัฒนาขอบเขตข้อมูลที่เหมาะสมที่สุด พบว่าสามารถใช้สร้างคำสำคัญได้สอดคล้อง และถูกต้องเช่นเดียวกับตัวอย่างที่ได้จากการตัดคำโดยอาศัยมนุษย์เป็นผู้กำหนด

งานวิจัย (Naenudorn and Niwattanakul, 2014) การค้นหาเว็บเซอร์วิสทำได้โดยการค้นหาจากยูดีดีไอซึ่งมีอยู่หลากหลาย ซึ่งส่วนใหญ่แล้วเป็นการค้นหาโดยใช้คำสำคัญ โดยไม่ได้คำนึงถึงความหมายของคำและไม่สามารถค้นหาเว็บเซอร์วิสในเชิงความหมายได้ และผลลัพธ์ที่ได้จากการค้นหาก็ไม่อาจทราบได้ว่าเว็บเซอร์วิสนั้นยังทำงานได้อยู่หรือไม่หรือมีคุณภาพเพียงใด ดังนั้นงานวิจัยนี้จึงได้ 1) รวบรวมเว็บเซอร์วิสจากหลายผู้ให้บริการโดยใช้เว็บเซอร์วิสที่เกี่ยวกับการพยากรณ์อากาศมาเป็นกรณีศึกษา 2) ทำการออกแบบฐานความรู้ของเว็บเซอร์วิสโดยใช้ออนโทโลยีและใช้ RAP-RDF API เพื่อสร้างส่วนติดต่อกับผู้ใช้สำหรับค้นหาเว็บเซอร์วิส 3) ทำการคัดเลือกเว็บเซอร์วิสตามคุณภาพการบริการ โดยมีปัจจัยสำหรับการคัดเลือก 4 ปัจจัย คือ เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ (Response Time) สภาพพร้อมใช้งาน (Availability) ความน่าเชื่อถือ (Reliability) และค่าใช้จ่ายในการเรียกใช้บริการ (Cost) จากผลการวิจัยทำให้สามารถค้นหาเว็บเซอร์วิสในเชิงความหมายได้และได้เว็บเซอร์วิสที่มีคุณภาพ

ตารางที่ 2.3 สรุปงานวิจัยที่เกี่ยวข้องด้านการคัดเลือกเว็บเซอร์วิส (Web Service Selection)

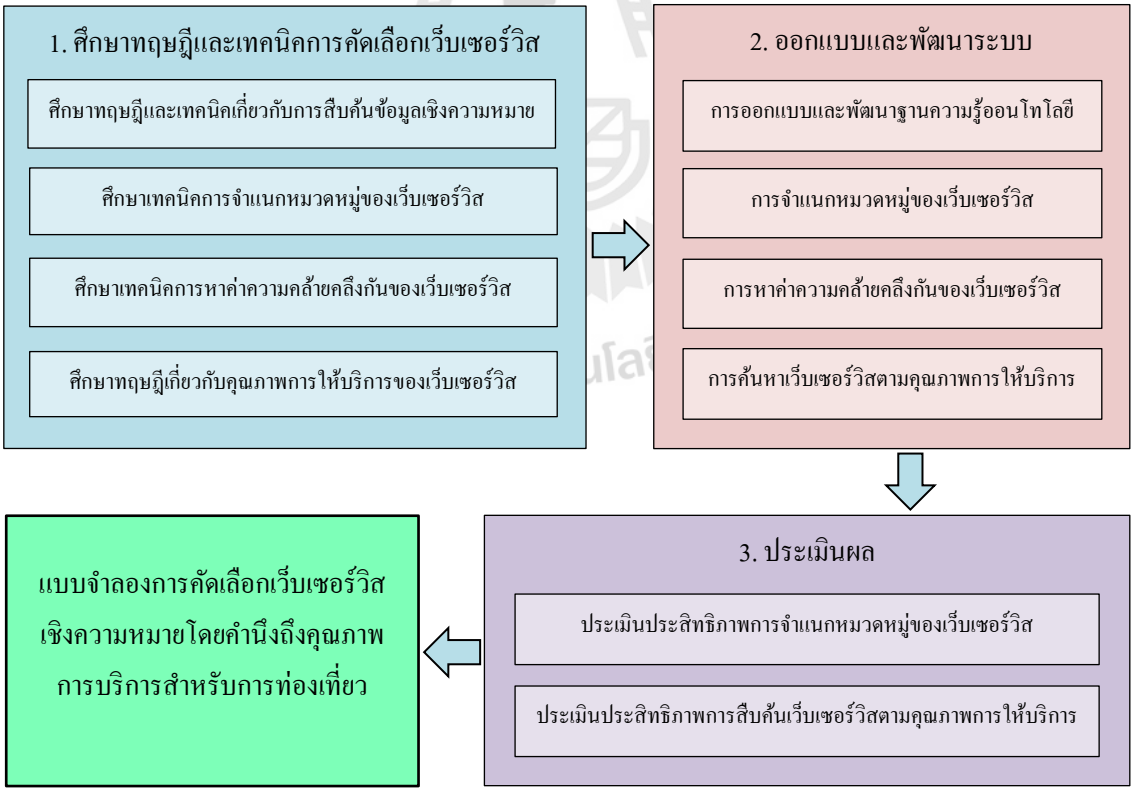
งานวิจัย / เทคนิควิธี	Similarity	Ontology	QoS	Ranking	User Preferences	Other
Ajao, T. A. and Deris, S., “Optimal Web Service Selection with Consideration for User's Preference”. (2013)	-	-	✓	-	✓	-
Badr,Y., et al.,“Enhancing Web Service Selection by User Preferences of Non-Functional Features”. (2008)	-	-	✓	-	✓	-
Zhao, L., Ren, Y., Li, M. and Sakurai, K., “Flexible service selection with user-specific QoS support in service oriented architecture”. (2012)	-	✓	✓	-	✓	-
Mohebi, A., Ding, C. and Chi, C-H., “Efficient QoS-based Service Selection with Consideration of User Requirements”. (2012)	-	-	✓	-	✓	-
Sodki, C., Youakim, B., and Frdrique, B. “Enhancing Web Service Selection by QoS Based Ontology and WS-Policy”. (2008)	-	✓	✓	✓	-	-
Chang, G., "QoS-Based Web Service Selection Approach". (2012)	-	-	✓	-	-	Genetic Algorithm
Benouaret, K., et al., "Majority-Rule-Based web service selection". (2012)	-	-	-	-	-	Majority Rule
Liu, Y., Ngu, A.H.H., and Zeng, L. “QoS Computation and Policing in Dynamic Web Service Selection”. (2004)	-	-	✓	✓	-	-
D'Mello, D.A., and V. S. Ananthanarayana. , “Dynamic selection mechanism for quality of service aware web services”. (2010)	-	-	✓	✓	-	-
Wang, H., et al., "Incomplete preference-driven web service selection". (2008)	-	-	-	✓	✓	CP-Nets
Wang, S., et al., "Multi-user web service selection based on multi-QoS prediction". (2013)	-	-	✓	-	-	Matching
งานวิจัยนี้	✓	✓	✓	✓	✓	WS-Classification, SKOS

จากตารางสรุปงานวิจัยที่เกี่ยวข้องด้านการคัดเลือกเว็บเซอร์วิสจะพบว่างานวิจัยที่ผู้วิจัยได้ทำนี้จะประกอบด้วยเทคนิคหรือวิธีการต่าง ๆ มากมายเพื่อทำให้แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวนี้มีประสิทธิภาพ เทคนิคหรือวิธีการทั้งหมด ได้แก่ การวัดความคล้ายคลึงกันของเว็บเซอร์วิส (Web Service Similarity) โดยใช้เทคนิคการวัดความคล้ายคลึงแบบโคไซน์ การสร้างออนโทโลยีของเว็บเซอร์วิส (Ontology Web Service) โดยประยุกต์จากออนโทโลยีของ Web Service Modeling Ontology (WSMO) การจัดการคลังคำ (Word Repository) โดยใช้สคอส (Simple Knowledge Organization System: SKOS) คุณภาพการให้บริการ (Quality of Service: QoS) การจัดลำดับข้อมูล (Ranking) การจำแนกหมวดหมู่ของเว็บเซอร์วิส (Web Service Classification) และผู้ใช้สามารถกำหนดคุณลักษณะต่าง ๆ ในระบบสืบค้นให้เหมาะสมกับความต้องการของผู้ใช้ได้ (User Preferences)

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 3
วิธีดำเนินการวิจัย

งานวิจัยนี้มีจุดมุ่งหมายเพื่อออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการ เนื้อหาในบทนี้ประกอบด้วยหัวข้อ วิธีวิจัย เครื่องมือที่ใช้ในการวิจัย และการวิเคราะห์ข้อมูล งานวิจัยนี้ได้ดำเนินการวิจัยตามกรอบแนวคิดการวิจัยใน 3 ขั้นตอน ได้แก่ ศึกษาทฤษฎีและเทคนิคต่าง ๆ เกี่ยวกับการคัดเลือกเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการ ออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว และประเมินผลแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว กรอบแนวคิดการวิจัยแสดงดังรูปที่ 3.1



รูปที่ 3.1 กรอบแนวคิดการวิจัย

3.1 วิธีวิจัย

การวิจัยนี้เป็นการวิจัยเชิงประยุกต์ เป็นการศึกษาวิจัยที่มีขั้นตอนการดำเนินงานที่ประยุกต์จากวงจรการพัฒนาระบบ (System Development Life Cycle: SDLC) มีรายละเอียดดังนี้

3.1.1 ศึกษาทฤษฎีและเทคนิคการคัดเลือกเว็บเซอร์วิส

3.1.1.1 การสืบค้นข้อมูลเชิงความหมาย

การสืบค้นเชิงความหมาย (Semantic Search) ซึ่งเป็นวิธีการค้นหาทรัพยากรสารสนเทศที่ต้องการจากคลังสารสนเทศหรือฐานข้อมูลภายใต้เงื่อนไขหรือกระบวนการคัดเลือกที่ผู้ใช้ต้องการ ทำให้สามารถเรียกข้อมูลที่ต้องการออกมาแสดงผลได้อย่างรวดเร็วและครอบคลุมถึงการค้นหาข้อมูลที่มีความเกี่ยวข้องและมีความสัมพันธ์เชิงความหมายด้วย การสืบค้นเชิงความหมายกลายมาเป็นแนวทางใหม่ที่จะช่วยให้การค้นหาข้อมูลมีประสิทธิภาพมากขึ้น ดังนั้นเพื่อให้การออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสนี้สามารถสืบค้นเว็บเซอร์วิสเชิงความหมายได้อย่างมีประสิทธิภาพ ผู้วิจัยจึงได้ศึกษาแนวคิดและทฤษฎีที่เกี่ยวข้องกับการสืบค้นเชิงความหมาย

3.1.1.2 การจำแนกหมวดหมู่ของเว็บเซอร์วิส

งานวิจัยนี้ใช้วิธีการทำเหมืองข้อมูล (Data Mining) ในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว งานวิจัยนี้ได้ทำการเปรียบเทียบว่าเทคนิคใดจะเหมาะสมกับข้อมูลเว็บเซอร์วิสที่ให้บริการที่เกี่ยวข้องด้านการท่องเที่ยวมากที่สุด และมีการประเมินผลว่าเทคนิคใดให้ความถูกต้องในการจำแนกหมวดหมู่ของเว็บเซอร์วิสได้ถูกต้องมากที่สุด ซึ่งเทคนิคต่าง ๆ ได้แก่ เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) เทคนิควิธีเค-เนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) และเทคนิควิธีเคมีนส์ (K-means)

3.1.1.3 การหาความคล้ายคลึงกันของเว็บเซอร์วิส

ศึกษาเทคนิคการหาความคล้ายคลึงกันของเว็บเซอร์วิสที่อยู่ในกลุ่มบริการเดียวกัน ใช้วิธีการหาค่าความคล้ายคลึงตามทฤษฎีเวกเตอร์สเปซโมเดล (Vector Space Model: VSM) คือการวัดความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) (Salton, 1988; Sidorov et al., 2014) โดยพิจารณาเปรียบเทียบความคล้ายคลึงกันจากชื่อของเว็บเซอร์วิส (Web Service Name) คำอธิบาย (Description) ชื่อโอเปอเรชัน (Operation Name) ข้อมูลเข้า (Input) และข้อมูลออก (Output)

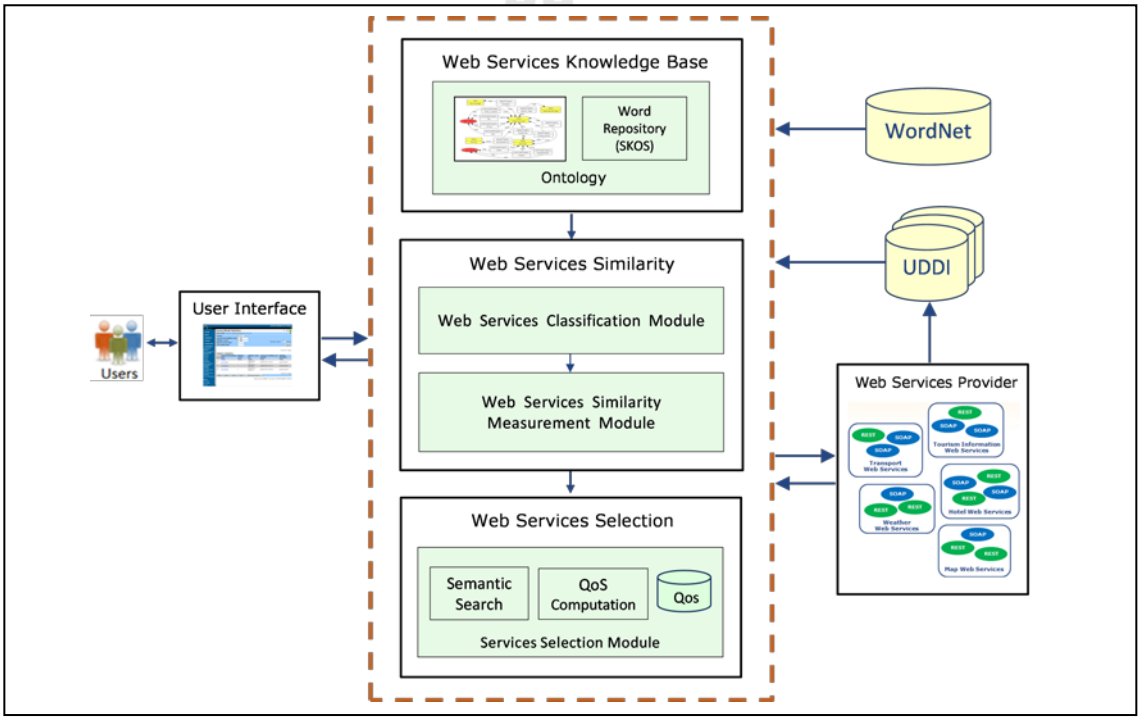
3.1.1.4 คุณภาพการให้บริการของเว็บเซอร์วิส

คุณภาพการให้บริการ (Quality of Service: QoS) เป็นเรื่องที่สำคัญสำหรับผู้ใช้บริการเว็บแอปพลิเคชันที่มีกระบวนการส่งข้อมูลสารสนเทศผ่านเครือข่ายอินเทอร์เน็ต เทคโนโลยีเว็บเซอร์วิสเป็นเทคโนโลยีหนึ่งที่ผู้ให้บริการนิยมนำมาใช้ในดำเนินงานทางธุรกิจเพื่อใช้

ในการนำเสนอสินค้าหรือบริการ ทำให้จำนวนเว็บเซอร์วิสบนอินเทอร์เน็ตมีจำนวนเพิ่มขึ้นอย่างรวดเร็ว เพราะผู้ให้บริการเว็บเซอร์วิส (Service Provider) มีการเสนอฟังก์ชันการบริการที่หลายหลายมากขึ้น (Li et.al, 2007) ดังนั้นเพื่อให้ผู้ใช้ได้ใช้บริการเว็บเซอร์วิสที่มีคุณภาพและประสิทธิภาพ เรื่องคุณภาพการให้บริการจึงเป็นปัจจัยที่สำคัญในการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวในงานวิจัยนี้

3.1.2 การออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว

3.1.2.1 การออกแบบสถาปัตยกรรมของระบบ SEMO หรือแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมาย (Semantic web services selection model หรือ SEMO อ่านว่า ซีโม่) แสดงได้ดังรูปที่ 3.2



รูปที่ 3.2 สถาปัตยกรรมของระบบ SEMO

สถาปัตยกรรมของระบบ SEMO ประกอบด้วย 7 องค์ประกอบ ได้แก่

1) WordNet (Miller, 1995) เวิร์ดเน็ตเป็นฐานข้อมูลคำศัพท์ภาษาอังกฤษ มีการจัดกลุ่มเป็นชุดของคำตามกลุ่มความหมาย โดยจะเชื่อมโยงโดยวิธีการของความสัมพันธ์ของแนวคิดและความหมายคำศัพท์ตามลำดับชั้น

2) UDDI หรือ Service Registry (UDDI-OASIS Standard, www, 2004) เป็นตัวกลางที่เก็บรวบรวมเว็บเซอร์วิสต่าง ๆ เพื่อให้ Service Requestor สามารถค้นหาเว็บเซอร์วิสต่าง ๆ ที่ Web Services Provider มาลงทะเบียนเพื่อบอกถึงรายละเอียดของการบริการเว็บเซอร์วิสของตนเองไว้

3) Web Services Provider เป็นผู้ให้บริการเว็บเซอร์วิสที่ได้ประกาศบริการของตนไว้บน UDDI

4) Web Services Knowledge Base เป็นขั้นตอนการสร้างฐานความรู้โดยการประยุกต์จากออนโทโลยีเว็บเซอร์วิส Web Service Modeling Ontology (WSMO) (W3C-WSMO, www, 2005) ของดับเบิ้ลยูทรีซี

5) Web Services Similarity เป็นขั้นตอนการหาค่าความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่ม ซึ่งก่อนการคำนวณหาค่าความคล้ายคลึง เว็บเซอร์วิสทั้งหมดจะต้องผ่านขั้นตอนการจำแนกหมวดหมู่ของเว็บเซอร์วิส

6) Web Services Selection เป็นขั้นตอนการออกแบบการสืบค้นเว็บเซอร์วิสเชิงความหมาย และการคัดเลือกเว็บเซอร์วิสที่มีคุณภาพ โดยมีการคำนวณหาค่าคะแนนคุณภาพของเว็บเซอร์วิส โดยพิจารณาจาก 5 ปัจจัย ได้แก่ เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ (Response Time) สภาพความพร้อมในการให้บริการ (Availability) ความเสถียรภาพหรือความน่าเชื่อถือ (Reliability) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) และความสำเร็จในการนำส่งข้อมูล (Throughput)

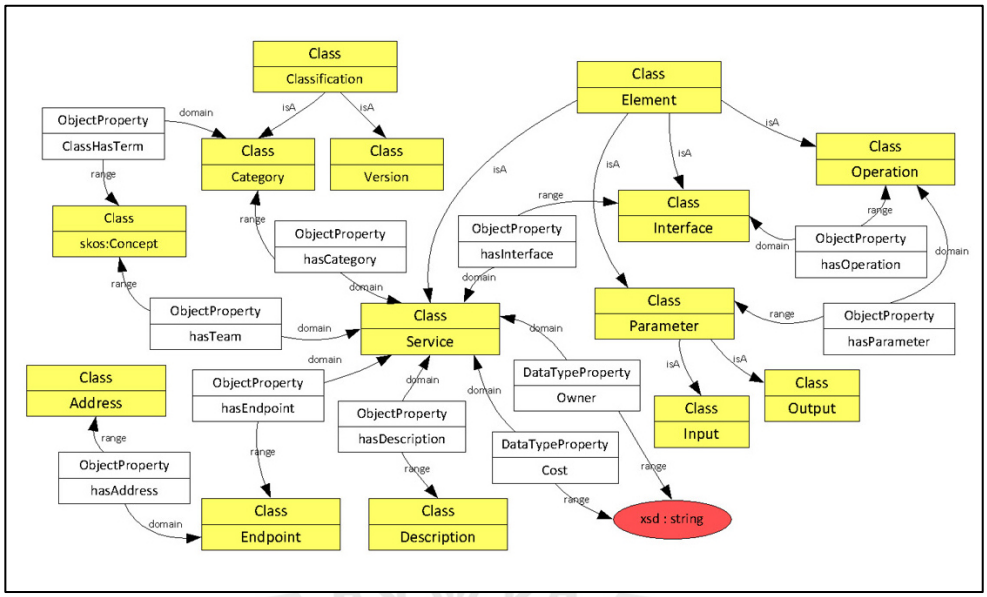
7) User Interface เป็นส่วนติดต่อกับผู้ใช้ เพื่อให้ผู้ใช้เข้าไปสืบค้นข้อมูลเว็บเซอร์วิสผ่านทางเว็บบราวเซอร์ จากนั้นระบบจะแสดงผลลัพธ์ที่ได้จากการค้นหา

3.1.2.2 การสร้างฐานความรู้โดยใช้ออนโทโลยี

งานวิจัยนี้มีการสร้างฐานความรู้ออนโทโลยีของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวโดยได้ประยุกต์จากออนโทโลยีเว็บเซอร์วิส Web Service Modeling Ontology (WSMO) (W3C-WSMO, www, 2005) ของดับเบิ้ลยูทรีซี ซึ่งมีขั้นตอนต่าง ๆ ดังนี้

1) เก็บรวบรวมข้อมูลเว็บเซอร์วิสจากยูดีดีไอรีจิสทรี (UDDI Registry) เช่น เว็บไซต์ http://www.xmethods.com, http://www.webservicex.net, http://www.service-repository.com, http://free-web-services.com, http://www.wsdream.net, http://www.odata.org, http://www.restfulwebservices.net และ http://www.programmableweb.com เป็นต้น จำนวนเว็บเซอร์วิสจากทั้งหมดประมาณ 200 เว็บเซอร์วิส ทำการคัดเลือกเฉพาะเว็บเซอร์วิสที่มีความสมบูรณ์ที่สุดจำนวน 76 เว็บเซอร์วิส เพื่อนำมาใช้ในการจำแนกหมวดหมู่ของเว็บเซอร์วิส จากนั้นรวมแอตทริบิวต์เข้าด้วยกัน ได้แก่ ชื่อของเว็บเซอร์วิส (Web Service Name) คำอธิบาย (Description) โอเปอเรชัน (Operation) ยูอาร์แอล (URL) ข้อมูลเข้า (Input) และข้อมูลออก (Output)

2) ออกแบบและพัฒนาออนโทโลยี ขั้นตอนนี้เป็นการออกแบบและพัฒนาออนโทโลยีสำหรับเว็บเซอร์วิสในโดเมนด้านการท่องเที่ยว โดยประยุกต์จากออนโทโลยีในโดเมนของเว็บเซอร์วิสในชื่อของ Web Service Modeling Ontology (WSMO) (W3C-WSMO, www, 2005) ของดับเบิ้ลยูทรีซี แล้วนำมาปรับปรุงเพื่อให้สอดคล้องกับบริบทของการท่องเที่ยว และจัดทำคลังคำโดยใช้สคอส (Simple Knowledge Organization System: SKOS)



รูปที่ 3.3 บางส่วนของออนโทโลยีเว็บเซอร์วิส

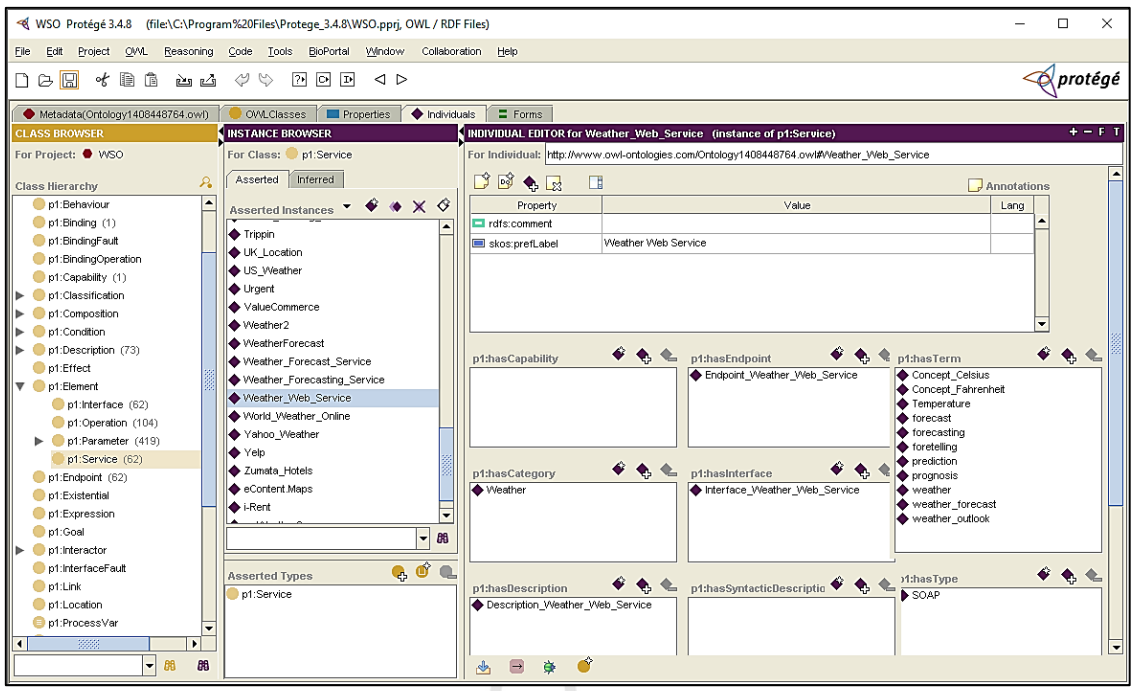
จากรูปที่ 3.3 ประกอบด้วย Class, Data Type Property และ Object Property รายละเอียดแสดงดังตารางที่ 3.1

ตารางที่ 3.1 รายละเอียดของออนโทโลยีเว็บเซอร์วิส

ที่	ชื่อ	ประเภท	คำอธิบาย
1	Element	Class	คลาสแทนข้อมูลอีลีเมนต์
2	Interface	Class	คลาสแทนข้อมูลอินเทอร์เฟส
3	Operation	Class	คลาสแทนข้อมูลโอเปอเรชันของการให้บริการเว็บเซอร์วิส
4	Parameter	Class	คลาสแทนข้อมูลพารามิเตอร์
5	Service	Class	คลาสแทนข้อมูลเว็บเซอร์วิส
6	Classification	Class	คลาสแทนข้อมูลประเภทของเว็บเซอร์วิส
7	Category	Class	คลาสแทนข้อมูลกลุ่มของเว็บเซอร์วิส

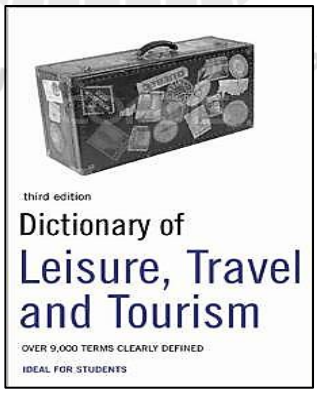
ตารางที่ 3.1 รายละเอียดของออนโทโลยีเว็บเซอร์วิส (ต่อ)

ที่	ชื่อ	ประเภท	คำอธิบาย
8	Version	Class	คลาสแทนข้อมูลเวอร์ชัน
9	Description	Class	คลาสแทนข้อมูลคำอธิบายของเว็บเซอร์วิส
10	Endpoint	Class	คลาสแทนข้อมูล Endpoint
11	Address	Class	คลาสแทนข้อมูลที่อยู่ของเว็บเซอร์วิส
12	skos:Concept	Class	คลาสแทนข้อมูล Term ต่าง ๆ
13	Input	Class	คลาสแทนข้อมูลนำเข้า
14	Output	Class	คลาสแทนข้อมูลผลลัพธ์ที่แสดงผล
15	Owner	Data type	เป็นการระบุค่าคงที่แสดงชื่อเจ้าของที่ให้บริการ
16	Cost	Data type	เป็นการระบุค่าคงที่แสดงค่าใช้จ่ายในการให้บริการ
17	hasInterface	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Service และ Class:Interface ว่าในแต่ละเซอร์วิสมีอินเทอร์เฟสใดบ้าง
18	hasOperation	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Interface และ Class:Operation ว่าในแต่ละอินเทอร์เฟสประกอบไปด้วยโอเปอเรชันใดบ้าง
19	hasParameter	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Operation และ Class:Parameter ว่าในแต่ละโอเปอเรชันประกอบด้วยพารามิเตอร์ใดบ้าง
20	hasCategory	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Service และ Class:Category ว่าในแต่ละเว็บเซอร์วิสจัดอยู่กลุ่มใด
21	hasDescription	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Service และ Class:Description ว่าในแต่ละเว็บเซอร์วิสมีคำอธิบายใด
22	hasEndpoint	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Service และ Class:Endpoint ว่าในแต่ละเว็บเซอร์วิสมี Endpoint ใด
23	hasAddress	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Endpoint และ Class:Address ว่าในแต่ละ Endpoint มีที่อยู่ใด
24	hasTerm	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Service และ Class:skos:Concept ว่าในแต่ละเว็บเซอร์วิสประกอบด้วยคำใดบ้างที่เกี่ยวข้อง
25	ClassHasTerm	Object	แสดงถึงความสัมพันธ์ระหว่าง Class:Category และ Class:skos:Concept ว่าเว็บเซอร์วิสในแต่ละกลุ่มประกอบด้วยคำใดบ้างที่เกี่ยวข้อง



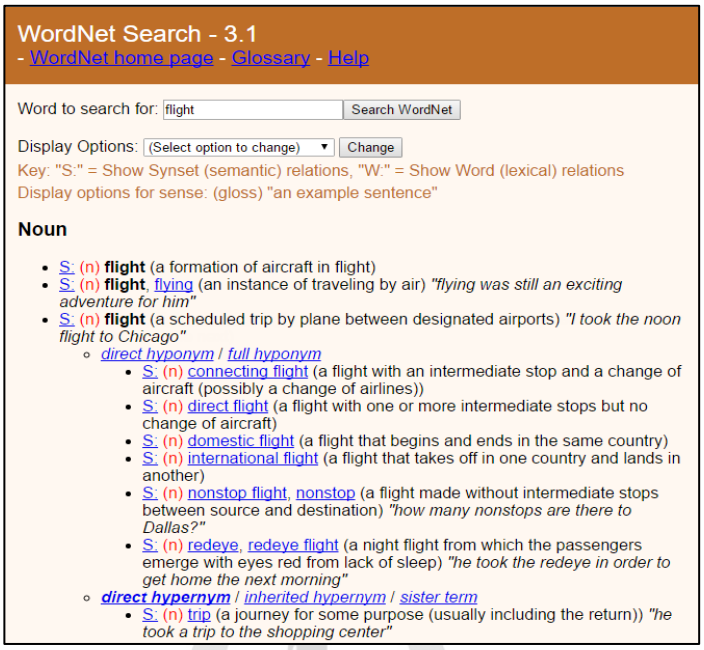
รูปที่ 3.4 แสดงออนโทโลยีเว็บเซอร์วิสในโปรแกรมโพรเทเจ

3) เก็บรวบรวมคำศัพท์ตัวอย่างที่เกี่ยวข้องกับการบริการด้านการท่องเที่ยวจากพจนานุกรมด้านการสันทนาการ การเดินทางและการท่องเที่ยว (Dictionary of Leisure, Travel and Tourism) (Russell and Roseby, 2006) จำนวน 200 คำ



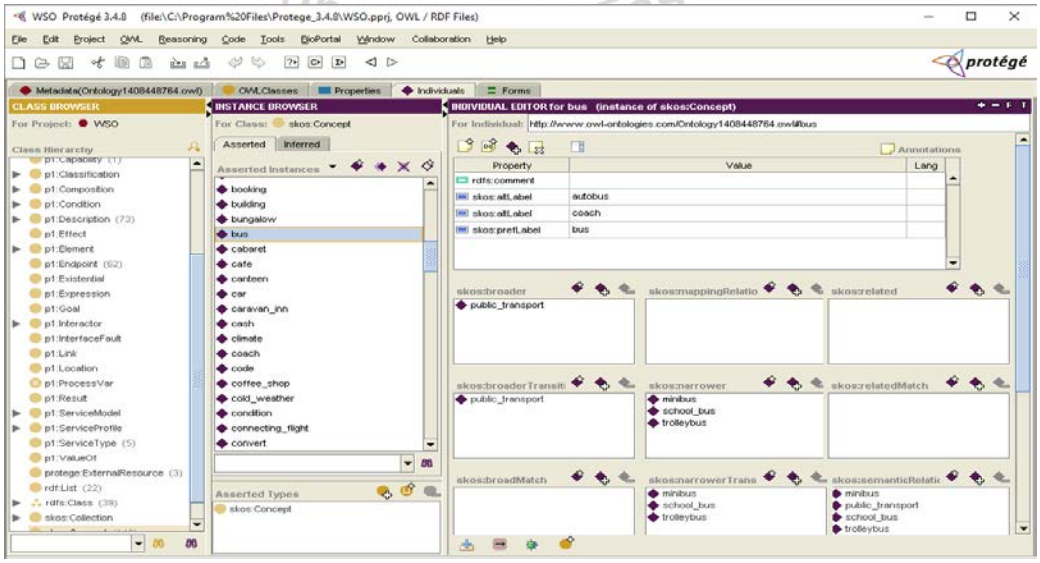
รูปที่ 3.5 พจนานุกรมด้านการสันทนาการ การเดินทางและการท่องเที่ยว (Russell and Roseby, 2006)

4) นำคำศัพท์ตัวอย่างที่ได้จาก Dictionary of Leisure, Travel and Tourism มาทำการค้นหาคำศัพท์ที่เกี่ยวข้องจากอภิธานศัพท์เวิร์ดเน็ต (WordNet) (Miller, 1995) ที่ประกอบไปด้วยคำศัพท์หลัก คำเหมือน คำที่กว้างกว่า คำที่แคบกว่า และคำที่เกี่ยวข้อง



รูปที่ 3.6 เว็บไซต์อภิธานศัพท์เวิร์ดเน็ตออนไลน์สำหรับค้นหาคำศัพท์ที่เกี่ยวข้อง (Miller, 1995)

5) บันทึกคำศัพท์ทั้งหมดที่ได้จากอภิธานศัพท์เวิร์ดเน็ตลงในคลังคำในออนโทโลยีที่ใช้สคอส (Simple Knowledge Organization System: SKOS) ในโปรแกรมโพรเทเจ (Protégé)



รูปที่ 3.7 ตัวอย่างคลังคำโดยใช้สคอสในโปรแกรมโพรเทเจ

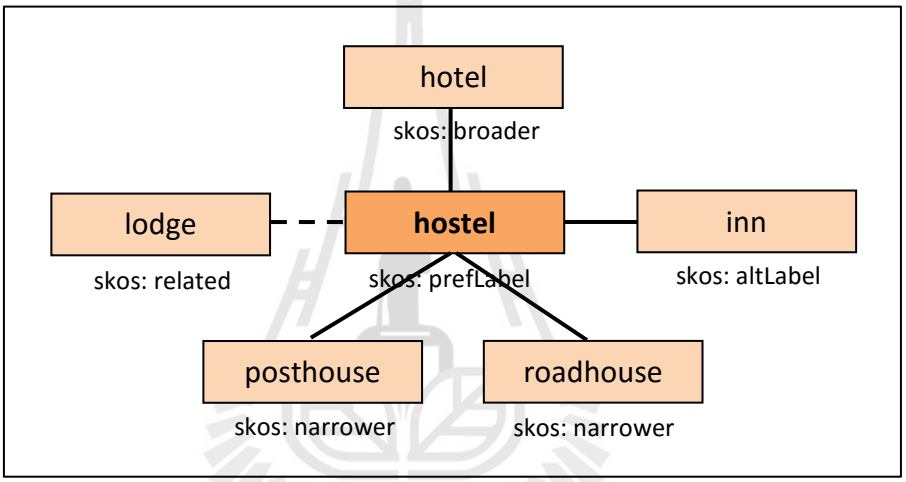
การจัดเก็บคำในคลังคำศัพท์ของงานวิจัยนี้ เป็นการจัดหมวดหมู่คำศัพท์โดยใช้สคอส (Simple Knowledge Organization System: SKOS) เป็นการอธิบายความสัมพันธ์เชิงความหมายของแนวคิดที่มีความสัมพันธ์อยู่ 3 ลักษณะคือ ความสัมพันธ์ในลักษณะเท่าเทียมกัน ความสัมพันธ์ในลักษณะลำดับชั้น และความสัมพันธ์ในลักษณะเกี่ยวข้องกันหรือความหมายคล้ายกัน โดยมีรูปแบบของการอธิบายคำศัพท์ที่เกี่ยวข้องกับคำหลัก ประกอบด้วย คำหลัก (Prefer Label) คำที่มีความหมายเหมือนกับคำหลัก (Alternate Label) คำที่มีความหมายกว้างกว่าคำหลัก (Broader) คำที่มีความหมายแคบกว่าคำหลัก (Narrower) และคำที่มีความหมายเกี่ยวข้องกับคำหลัก (Related) ตัวอย่างคำศัพท์ในคลังคำในฐานความรู้ออนโทโลยีแสดงในตารางที่ 3.2

ตารางที่ 3.2 ตัวอย่างคำศัพท์บางส่วนในคลังคำในฐานความรู้ออนโทโลยี

คำหลัก (prefLabel)	คำเหมือน (altLabel)	คำที่กว้างกว่า (broader)	คำที่แคบกว่า (narrower)	คำที่เกี่ยวข้อง (related)
flight		trip	direct flight, connecting flight, domestic_flight, international flight	
forecast	prognosis	prediction, foretelling, forecasting, prognostication	weather forecast, weather outlook, financial forecast	prospect
hotel		building, edifice	hostel, auberge, resort, resort hotel, holiday resort, spa	
hostel	inn	hotel	posthouse, roadhouse	lodge
house	home	building, edifice	bungalow, cottage, guesthouse, mansion, villa, icehouse	

ตารางที่ 3.2 ตัวอย่างคำศัพท์บางส่วนในคลังคำในฐานความรู้ออนโทโลยี (ต่อ)

คำหลัก (prefLabel)	คำเหมือน (altLabel)	คำที่กว้างกว่า (broader)	คำที่แคบกว่า (narrower)	คำที่เกี่ยวข้อง (related)
public transport		transport	bus, express, shuttle, train	transportation, transit
restaurant	eatery, eating house, eating place	building, edifice	cafe, canteen, coffee shop, diner, teashop	



รูปที่ 3.8 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “hostel”

จากรูปที่ 3.8 แสดงให้เห็นถึงโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลักคำว่า “hostel” ซึ่งมีคำว่า “inn” เป็นคำที่มีความหมายเหมือนกัน มีคำที่มีความหมายกว้างกว่าคือ “hotel” คำที่มีความหมายที่เกี่ยวข้องคือ “lodge” และคำที่มีความหมายแคบกว่าคือ “posthouse” และ “roadhouse”

3.1.2.3 การจำแนกหมวดหมู่ของเว็บเซอร์วิส (Web services Classification)

กระบวนการนี้เป็นขั้นตอนการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว ซึ่งกระบวนการในการจำแนกหมวดหมู่ของเว็บเซอร์วิสมี 7 ขั้นตอนดังต่อไปนี้

1) เขียนโปรแกรมภาษาพีเอชพี (PHP) ร่วมกับแรพเอพีไอ (RAP API) เพื่อค้นหาข้อมูลเว็บเซอร์วิสจากออนโทโลยีโดยใช้ภาษาสอบถามข้อมูลสปาร์เคิล (SPARQL) ในการสืบค้นข้อมูลในออนโทโลยีที่ถูกส่งออกมาในรูปแบบของภาษาโอดับเบิ้ลยูแอล (OWL) ข้อมูลเว็บ

เซอร์วิสที่ต้องการได้แก่ ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบายเว็บเซอร์วิส (Description) ชื่อโอเปอเรชัน (Operation) ข้อมูลเข้า (Input) ข้อมูลออก (Output) หมายเลขของเว็บเซอร์วิส (Web Service ID) และคำที่เกี่ยวข้องกับเว็บเซอร์วิส (Term) ตัวอย่างคำสั่งภาษาสปาร์เคิล (SPARQL) และผลลัพธ์ แสดงดังรูปที่ 3.9 และรูปที่ 3.10

ตัวอย่างคำสั่งภาษาสปาร์เคิล (SPARQL) สอบถามข้อมูลในออนโทโลยี ในการค้นข้อมูลชื่อเว็บเซอร์วิส และคำอธิบายของเว็บเซอร์วิส โดยเรียงลำดับจากชื่อเว็บเซอร์วิส

```
SELECT ?wsName ?description
WHERE { ?wsID p1:hasDescription ?desID .
?wsID skos:prefLabel ?wsName .
?desID rdfs:comment ?description
}
ORDER BY ASC(?wsName)
```

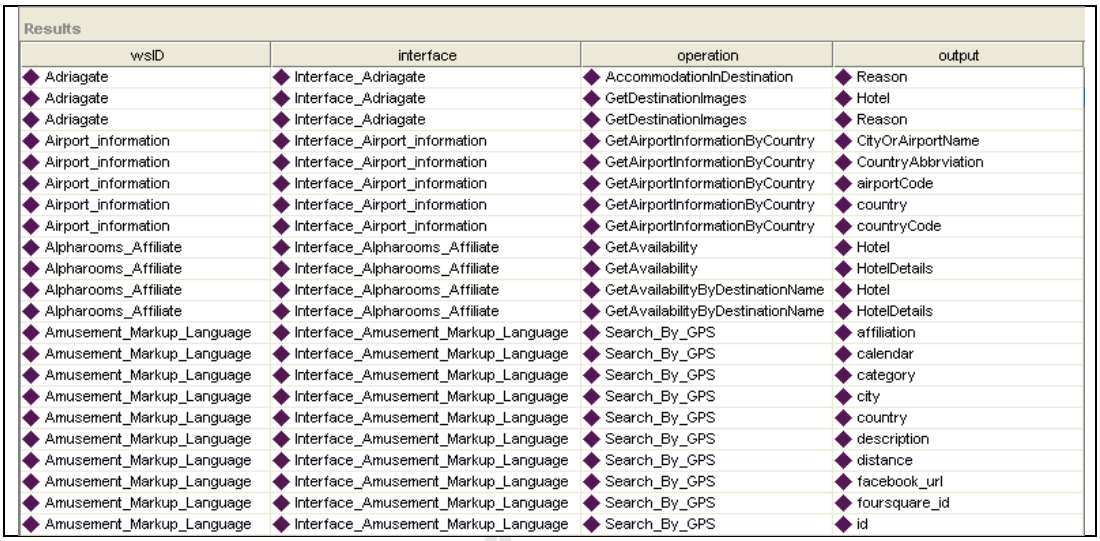
Results

wsName	description
Adriagate	Adriagate is a travel service specializing in holiday accommodations for Croatia. Their online offerings rep...
Airport information	Get Airport Code, CityOrAirport Name, Country, Country Abbrv, CountryCode,GMT Offset Runway Lengtl...
Alpharooms Affiliate	The Alpharooms Affiliate website helps travelers find discount hotels, budget flights, and cheap holidays...
Amusement Markup Language	Amusement Markup Language API gathers information on attractions, events and points of interest aroun...
Australian PostCode	Australian Postcode ,Location Web service Get Australian Postcode by partial / full Location Name Get Au...
British Airways	British Airways (BA) is a full service global airline with an extensive global route network based out of th...
Checkfront	Checkfront is a booking service that allows businesses to manage their schedule based inventories, cen...
Cleartrip Hotel	Cleartrip provides booking for flights, hotels and trains, as well as other travel services across the world...
Close To The Point	Close To The Point (CTTP) provides geolocation data so that app providers can better interact with their u...
DOTS Fast Weather	Weather forecasting service
EasyToBook	Easytobook will release a REST API. Easytobook provides hotel bookings for more than 120,000 hotels in...
Expedia	Expedia is a world leader in travel and technology, and the EAN Developer Hub gives developers free ac...
Five Day Weather	The 5 Day Weather API accepts a city as a single parameter and returns a 5 day forecast. Specifically, th...
FlitWays	The FlitWays online booking engine and app allows travelers to pre-book a ride and choose a preferred r...
Foursquare Real-time	Foursquare is a local search and discovery service company that provides customers with information o...
GTA Travel	GTA-Travel is a global travel product provider that has sourced a broad portfolio of accommodation optio...

รูปที่ 3.9 ผลลัพธ์จากการสอบถามข้อมูล (Query) ชื่อของเว็บเซอร์วิสและคำอธิบาย

ตัวอย่างคำสั่งภาษาสปาร์เคิล (SPARQL) สอบถามข้อมูลในออนโทโลยี ในการค้นข้อมูลชื่อโอเปอเรเตอร์และข้อมูลเอาท์พุทของเว็บเซอร์วิส โดยเรียงลำดับจาก ID ของเว็บเซอร์วิส

```
SELECT ?wsID ?interface ?operation ?output
WHERE { ?wsID p1:hasInterface ?interface .
?interface p1:hasOperation ?operation .
?operation p1:hasOutput ?output
}
ORDER BY ASC(?wsID)
```



Results

wsID	interface	operation	output
Adriagate	Interface_Adriagate	AccommodationInDestination	Reason
Adriagate	Interface_Adriagate	GetDestinationImages	Hotel
Adriagate	Interface_Adriagate	GetDestinationImages	Reason
Airport_information	Interface_Airport_information	GetAirportInformationByCountry	CityOrAirportName
Airport_information	Interface_Airport_information	GetAirportInformationByCountry	CountryAbbrviation
Airport_information	Interface_Airport_information	GetAirportInformationByCountry	airportCode
Airport_information	Interface_Airport_information	GetAirportInformationByCountry	country
Airport_information	Interface_Airport_information	GetAirportInformationByCountry	countryCode
Alpharooms_Affiliate	Interface_Alpharooms_Affiliate	GetAvailability	Hotel
Alpharooms_Affiliate	Interface_Alpharooms_Affiliate	GetAvailability	HotelDetails
Alpharooms_Affiliate	Interface_Alpharooms_Affiliate	GetAvailabilityByDestinationName	Hotel
Alpharooms_Affiliate	Interface_Alpharooms_Affiliate	GetAvailabilityByDestinationName	HotelDetails
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	affiliation
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	calendar
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	category
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	city
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	country
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	description
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	distance
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	facebook_url
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	foursquare_id
Amusement_Markup_Language	Interface_Amusement_Markup_Language	Search_By_GPS	id

รูปที่ 3.10 ผลลัพธ์จากการสอบถามข้อมูล Interface, Operation และ Output ของเว็บเซอร์วิส

2) ขั้นตอนการตัดเครื่องหมาย สัญลักษณ์ ตัวเลข และช่องว่าง ออกจากชุดสตริงหลัก หลังจากการสอบถามข้อมูลในขั้นตอนที่ 1 จะได้ข้อมูลเว็บเซอร์วิสที่ต้องการได้แก่ ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบายเว็บเซอร์วิส (Description) ชื่อโอเปอเรชัน (Operation) ข้อมูลเข้า (Input) ข้อมูลออก (Output) หมายเลขของเว็บเซอร์วิส (Web Service ID) และคำที่เกี่ยวข้องกับเว็บเซอร์วิส (Term) ซึ่งข้อมูลทั้งหมดที่ได้จะเป็นชุดสตริงยาว ๆ ที่อาจมีเครื่องหมาย สัญลักษณ์ ตัวเลข หรือช่องว่าง รวมอยู่ด้วย เช่น ! @ # $ % ^ & * ( ) - + = | { } [ ] : / \ . ? , [space] หรือตัวเลข (Numeric) เป็นต้น ดังนั้นจึงจำเป็นต้องทำการทำความสะอาดข้อมูล (Data Cleaning) เพื่อกำจัดข้อมูลที่ไม่เกี่ยวข้องออกไป ดังตัวอย่างในตารางที่ 3.3

ตารางที่ 3.3 ตัวอย่างผลลัพธ์ที่ได้จากการทำความสะอาดข้อมูล (Data Cleaning) ให้กับชุดสตริง

ชุดสตริงเดิม	ชุดสตริงใหม่
British Airways (BA) is a full service global airline with an extensive global route network based out of the United Kingdom. British Airways offers developers a REST API for the integration of BA data into third party applications. Current public API offerings include: Flight Time Info, for the latest flight information distributed by BA; Flight Offer, for the latest BA flight offers; and Availability, for hotel, car hire, and package offer information distributed by BA. An API key is required. Output: #flightID, flightNumber [id]	British Airways BA is a full service global airline with an extensive global route network based out of the United Kingdom British Airways offers developers a REST API for the integration of BA data into third party applications Current public API offerings include Flight Time Info for the latest flight information distributed by BA Flight Offer for the latest BA flight offers and Availability for hotel car hire and package offer information distributed by BA An API key is required Output flightID flightNumber id

3) การตัดแบ่งคำออกจากสตริง (Word Segmentation) หลังจากการตัดข้อมูลที่ไม่จำเป็นออกในขั้นตอนที่ 2 จะได้ข้อมูลเว็บเซอร์วิสที่สะอาด การตัดแบ่งคำนี้จะรวมถึงการแยกตัวพิมพ์ใหญ่ที่เป็นคำย่อออกด้วย เช่น USA BA 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) การตัดคำหยุด (Stop words) เป็นขั้นตอนตัดคำที่ไม่จำเป็นออก ซึ่งคำหยุดหรือ Stop words เป็นคำขยายให้แก่คำอื่น ๆ แต่ไม่มีความหมายในตัวเอง เมื่อตัดออกแล้วไม่ทำให้ใจความของเอกสารเปลี่ยนแปลง เช่น คำบุพบทเป็นคำที่ใช้เชื่อมคำหรือกลุ่มคำให้สัมพันธ์กัน คำสรรพนามเป็นคำที่ใช้แทนคำนามที่กล่าวถึงมาแล้วในประโยค คำสันธานเป็นคำที่ทำหน้าที่เชื่อมคำกับคำ เป็นต้น คำหยุดมักเป็นคำที่ปรากฏบ่อยครั้งและมีอยู่ในทุกเว็บเซอร์วิส จึงถือได้ว่าคำหยุดเป็นคุณลักษณะที่ไม่เกี่ยวข้องในการค้นคืนหรือการจำแนกหมวดหมู่ ดังนั้นจึงควรกำจัดคำหยุดซึ่งจะช่วยประหยัดทั้งพื้นที่และเวลาในการประมวลผล ตัวอย่างคำหยุด เช่น a, an, and, at, been, both, but, by, can, cannot, do, done, for, full, go, give, in, is, of, on, or, other, out, than, the, to, what, while, who, why, will หรือ with เป็นต้น ตัวอย่างผลลัพธ์ที่ได้จากการตัดคำหยุดแสดงดังตารางที่ 3.4 ซึ่งชุดของคำศัพท์ก่อนตัดคำหยุดมีจำนวน 21 คำ หลังจากตัดคำหยุดเหลือเพียง 13 คำ

ตารางที่ 3.4 ตัวอย่างผลลัพธ์ที่ได้จากการตัดคำหยุด (Stop words)

ชุดคำศัพท์ก่อนตัดคำหยุด	ชุดคำศัพท์หลังตัดคำหยุด
British | Airways | BA | is | a | full | service | global | airline | with | an | extensive | global | route | network | based | out | of | the | United | Kingdom	British | Airways | BA | service | global | airline | extensive | global | route | network | based | United | Kingdom

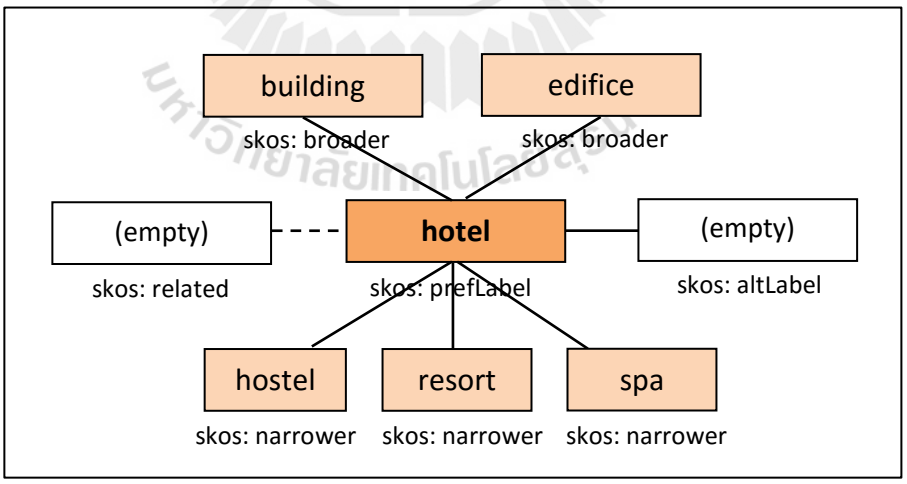
5) การหารากศัพท์ (Stemming Words) ขั้นตอนนี้เป็นกระบวนการในการหารูปเดิมของคำเพื่อปรับรวมให้เป็นคำเดียวกัน การหารากศัพท์ทำให้สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพในการจำแนกหมวดหมู่ของเว็บเซอร์วิสได้ เช่น รากศัพท์ของคำว่า “Booking” คือคำว่า “Book” รากศัพท์ของคำว่า “Transportation” คือคำว่า “Transport” รากศัพท์ของคำว่า “Members” คือคำว่า “Member” เป็นต้น ตัวอย่างผลลัพธ์ที่ได้จากการหารากศัพท์แสดงดังตารางที่ 3.5

ตารางที่ 3.5 ตัวอย่างผลลัพธ์ที่ได้จากการหารากศัพท์ (Stemming Words)

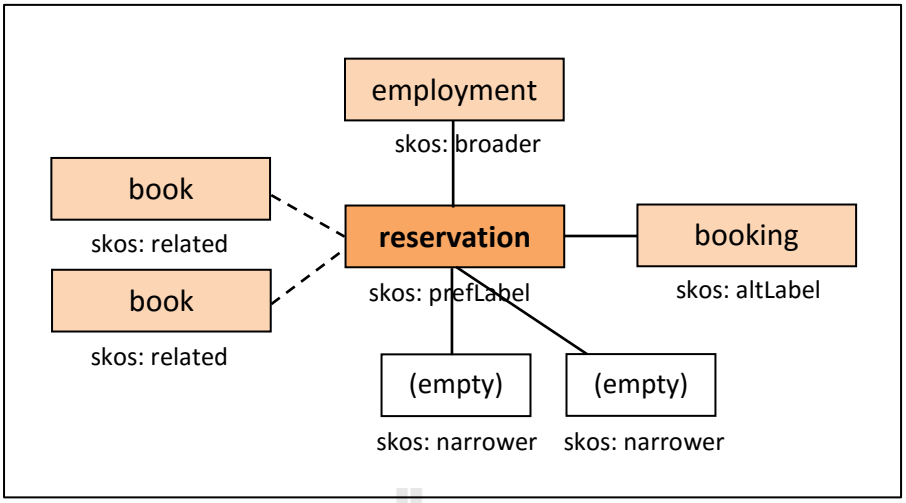
ชุดคำศัพท์ก่อนหารากศัพท์	ชุดคำศัพท์หลังหารากศัพท์
British | Airways | BA | service | global | airline | extensive | global | route | network | based | United | Kingdom	British | Airwai | BA | servic | global | airlin | extens | global | rout | network | base | Unite | Kingdom

6) การเปรียบเทียบระหว่างคำศัพท์กับคลังคำในออนโทโลยีเพื่อหาคำอื่นที่เกี่ยวข้อง ขั้นตอนนี้จะนำคำศัพท์แต่ละคำไปเปรียบเทียบกับคลังคำในออนโทโลยีที่จัดเก็บในรูปแบบสคอส (SKOS) เพื่อหาคำหลัก (Prefer Label) คำที่มีความหมายเหมือนกับคำหลัก (Alternate Label) คำที่มีความหมายกว้างกว่า (Broader) และคำที่มีความหมายแคบกว่า (Narrower) จากนั้นนำคำที่ได้ทั้งหมดแทนค่าเพิ่มเข้าไปในชุดสตริงหลัก

ตัวอย่างนี้นำคำหลักจำนวน 2 คำ คือ “hotel” และ “reservation” ไปเปรียบเทียบกับคลังคำในออนโทโลยีเพื่อหาคำอื่นที่เกี่ยวข้อง คำหลัก “hotel” ได้คำที่เกี่ยวข้องได้แก่ “building” “edifice” “hostel” “resort” และ “spa” ส่วนคำหลัก “reservation” ได้คำที่เกี่ยวข้องได้แก่ “booking” “employment” “reserve” และ “book” ดังแสดงในรูปที่ 3.11 และรูปที่ 3.12 และตัวอย่างผลลัพธ์การแทนค่าเข้าไปในชุดสตริงหลักแสดงดังตารางที่ 3.6



รูปที่ 3.11 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “hotel”



รูปที่ 3.12 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “reservation”

ตารางที่ 3.6 ตัวอย่างผลลัพธ์ที่ได้จากการเปรียบเทียบระหว่างคำศัพท์กับคลังคำในออนโทโลยี

ชุดคำศัพท์ในสตริงหลัก	ผลลัพธ์ที่ได้
hotel reservation	**hotel** building edifice hostel resort spa **reservation** booking employment reserve book

7) เข้าสู่กระบวนการทำเหมืองข้อมูล (Data Mining) เพื่อจำแนกหมวดหมู่ของเว็บเซอร์วิส หลังจากผ่านกระบวนการเตรียมข้อมูล (Data Preparation) การทำความสะอาดข้อมูล (Data Cleaning) และขั้นตอนตามหลักการเชิงความหมายแล้ว ขั้นตอนนี้เป็นกระบวนการทำเหมืองข้อมูลเพื่อจำแนกหมวดหมู่ของเว็บเซอร์วิส 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 คือชุดข้อมูลของเว็บเซอร์วิสที่ผ่านการหาความสัมพันธ์ของคำเชิงความหมายแล้ว ประกอบด้วยแอตทริบิวต์ต่าง ๆ จำนวน 12 แอตทริบิวต์ ได้แก่ รหัสเว็บเซอร์วิส (Web Service ID) ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบาย (Description) ชื่อกลุ่ม (Group Name) ชื่อโอเปอเรชัน (Operation Name) ข้อมูลเข้า (Input) ข้อมูลออก (Output) คำศัพท์หลัก (Terms) คำเหมือน (Alternate Terms) คำที่กว้างกว่า (Broader

Terms) คำที่แคบกว่า (Narrower Terms) และคำที่เกี่ยวข้อง (Releted Terms) จากนั้นใช้เทคนิควิธีต่าง ๆ ในการทดลอง เปรียบเทียบด้วย 5 เทคนิควิธี ได้แก่

ก) เทคนิควิธีเคมีนส์ (K-means) โดยกำหนดค่า k เท่ากับ 5 (เว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวมีทั้งหมด 5 กลุ่ม)

ข) เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) โดยกำหนดค่า cross-validation เท่ากับ 10

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

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

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

(หมายเหตุ: เทคนิควิธี ข. ค. ง. และ จ. ผู้วิจัยได้ทำการทดลองนอกรอบโดยกำหนดค่า cross-validation ระหว่าง 3 ถึง 20 พบว่าค่า cross-validation เท่ากับ 10 มีความเหมาะสมที่สุด เนื่องจากให้ค่าความถูกต้องในการจำแนกหมวดหมู่ของเว็บเซอร์วิสเฉลี่ยสูงที่สุด)

3.1.2.4 การคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส (Web services Similarity)

หลังจากเสร็จขั้นตอนการจำแนกหมวดหมู่ของเว็บเซอร์วิสแล้ว ขั้นตอนนี้จะทำการคำนวณเพื่อหาค่าความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่ม โดยใช้วิธีการหาค่าความคล้ายคลึงตามทฤษฎี Vector Space Model (VSM) คือการวัดความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) (Salton, 1988; Sidorov et al., 2014) โดยพิจารณาเปรียบเทียบความคล้ายคลึงกันจากชื่อของเว็บเซอร์วิส (Web Service Name) ชื่อโอเปอเรชัน (Operation Name) คำอธิบาย (Description) ข้อมูลเข้า (Input) และข้อมูลออก (Output)

ขั้นตอนการคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิสมีดังนี้

1) เขียนโปรแกรมภาษาพีเอชพี โดยเขียนร่วมกับเอพีไอแรพ เพื่อดึงข้อมูลเว็บเซอร์วิสจากออนโทโลยีโดยใช้ภาษาสอบถามข้อมูลสปาร์เคิล (SPARQL) ในการสืบค้นข้อมูลในออนโทโลยีที่ถูกส่งออกมาในรูปแบบของภาษาโอดับเบิ้ลยูแอล (OWL) ข้อมูลเว็บเซอร์วิสที่ต้องการได้แก่ ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบายเว็บเซอร์วิส (Description) ชื่อโอเปอเรชัน (Operation) ข้อมูลเข้า (Input) ข้อมูลออก (Output) และข้อมูลกลุ่มของเว็บเซอร์ซิส (Category)

2) การตัดเครื่องหมาย สัญลักษณ์ ตัวเลข และช่องว่าง ออกจากชุดสตริงหลัก เช่น ! @ # $ % ^ & * ( ) - + = | { } [ ] : / \ . ? , [space] หรือตัวเลข (Numeric) เป็นต้น ดังนั้นจึงจำเป็นต้องทำการทำความสะอาดข้อมูล (Data Cleaning) เพื่อกำจัดข้อมูลที่ไม่เกี่ยวข้องออกไป

3) การตัดแบ่งคำออกจากสตริง (Word Segmentation) หลังจากการตัดข้อมูลที่ไม่จำเป็นออกในขั้นตอนที่ 2 จะได้ข้อมูลเว็บเซอร์วิสที่สะอาด ขั้นตอนนี้จะได้ผลลัพธ์ที่เป็นชุดของคำศัพท์

4) การตัดคำหยุด (Stop words) เป็นขั้นตอนตัดคำที่ไม่จำเป็นออก ซึ่งคำหยุดหรือ Stop words เป็นคำขยายให้แก่คำอื่น ๆ แต่ไม่มีความหมายในตัวเอง เมื่อตัดออกแล้วไม่ทำให้ใจความของเอกสารเปลี่ยนแปลง ดังนั้นจึงควรกำจัดคำหยุดซึ่งจะช่วยประหยัดทั้งพื้นที่และเวลาในการประมวลผล ตัวอย่างคำหยุด เช่น a, an, and, at, been, both, but, by, can, cannot, do, done, for, full, go, give, in, is, of, on, or, other, out, than, the, to, what, while, who, why, will หรือ with เป็นต้น

5) กำหนดค่าน้ำหนักให้กับแต่ละองค์ประกอบของเว็บเซอร์วิสตามความสำคัญ ได้แก่ ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบายเว็บเซอร์วิส (Description) ชื่อโอเปอเรชัน (Operation) ข้อมูลเข้า (Input) และข้อมูลออก (Output)

6) หารายการคำจากชุดคำที่ได้ของเว็บเซอร์วิสทั้งหมด และตัดคำที่ซ้ำกันออกจากนั้นรวมเข้ากันเป็นสตริงชุดใหญ่

7) คำนวณค่าความถี่ของคำ (หาค่า F) นับความถี่คำตามรายการคำที่ได้มาโดยแยกแต่ละเว็บเซอร์วิส และแยกแต่ละองค์ประกอบของเว็บเซอร์วิส (นับ F ตามเว็บเซอร์วิส)

8) คำนวณค่าความถี่ของคำ (หาค่า F) นับความถี่คำตามรายการคำที่ได้มาโดยแยกแต่ละกลุ่มเว็บเซอร์วิส (Category) และแยกแต่ละองค์ประกอบ (นับ F ตามกลุ่ม)

9) คำนวณค่า TF จากผลลัพธ์ในข้อที่ 7 ตามสมการที่ 3.1

$$\mathrm{tf}(t,d) = 0.5 + 0.5 \cdot \frac{f_{t,d}}{\max\{f_{t',d} : t' \in d\}} \tag{3.1}$$

10) คำนวณค่า IDF จากผลลัพธ์ในข้อที่ 8 ตามสมการที่ 3.2

$$\mathrm{idf}(t,D) = \log \frac{N}{|\{d \in D : t \in d\}|} \tag{3.2}$$

11) คำนวณค่า TF-IDF จากผลลัพธ์ในข้อที่ 9 และข้อที่ 10 ตามสมการที่ 3.3

$$\mathrm{tfidf}(t,d,D) = \mathrm{tf}(t,d) * \mathrm{idf}(t,D) \tag{3.3}$$

12) รวมคะแนนทุกองค์ประกอบของเว็บเซอร์วิสโดยการถ่วงน้ำหนักตามสมการที่ 3.4

$$tfidf_{ws}(t,d) = \sum_{obj} \left(tfidf_{obj}(t,d) * weigth_{obj} \right) \quad (3.4)$$

โดยที่ $tfidf_{ws}(t,d)$ คือ ค่าถ่วงน้ำหนัก TF-IDF ของแต่ละเว็บเซอร์วิส

t คือ คำแต่ละคำ

d คือ กลุ่มของเว็บเซอร์วิส (Category)

$tfidf_{obj}(t,d)$ คือ ค่า TF-IDF ของแต่ละเว็บเซอร์วิสของ องค์ประกอบ ได้แก่ name, description, operation, input และ output

13) นำค่า TF-IDF ของเว็บเซอร์วิสที่จะเปรียบเทียบกัน มาคำนวณตามสมการโคไซน์ (Cosine Similarity) ดังสมการที่ 3.5

$$\text{similarity} = \frac{\text{A} \cdot \text{B}}{||\text{A}||\,||\text{B}||} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \quad (3.5)$$

14) ปรับปรุงค่าคะแนนให้อยู่ในรูปแบบที่เหมาะสม (Normalize) มีค่าที่เป็นไปได้อยู่ระหว่าง 0.5-1 คะแนน ตามสมการที่ 3.6

$$normalize = \frac{(sim+max-2min)}{(2max-2min)} \quad (3.6)$$

โดยที่

$normalize$ คือ ค่าคะแนนความคล้ายคลึงที่ปรับปรุงแล้ว

sim คือ ค่าความค่าคล้ายคลึงเว็บเซอร์วิสแต่ละตัว

max คือ ค่าสูงสุดของค่าความคล้ายคลึงของเว็บเซอร์วิสทั้งหมด

min คือ ค่าต่ำสุดของค่าความคล้ายคลึงของเว็บเซอร์วิสทั้งหมด

ตัวอย่างการคำนวณค่าความคล้ายคลึงกันของเว็บเซอร์วิส ตัวอย่างนี้มีเว็บเซอร์วิสจำนวน 4 เว็บเซอร์วิส เว็บเซอร์วิสแต่ละตัวประกอบไปด้วยข้อมูลเข้า (Input) และข้อมูลออก (Output) อย่างละ 2 ตัว ถ้าเว็บเซอร์วิสใดมีอินพุทหรือเอาท์พุทที่คล้ายกันจะแทนค่าเวกเตอร์ของคำนั้นเท่ากับ 1 และถ้าไม่คล้ายกันจะแทนค่าเวกเตอร์ของคำนั้นเท่ากับ 0

ตารางที่ 3.7 การแทนค่าเพื่อคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส รอบที่ 1

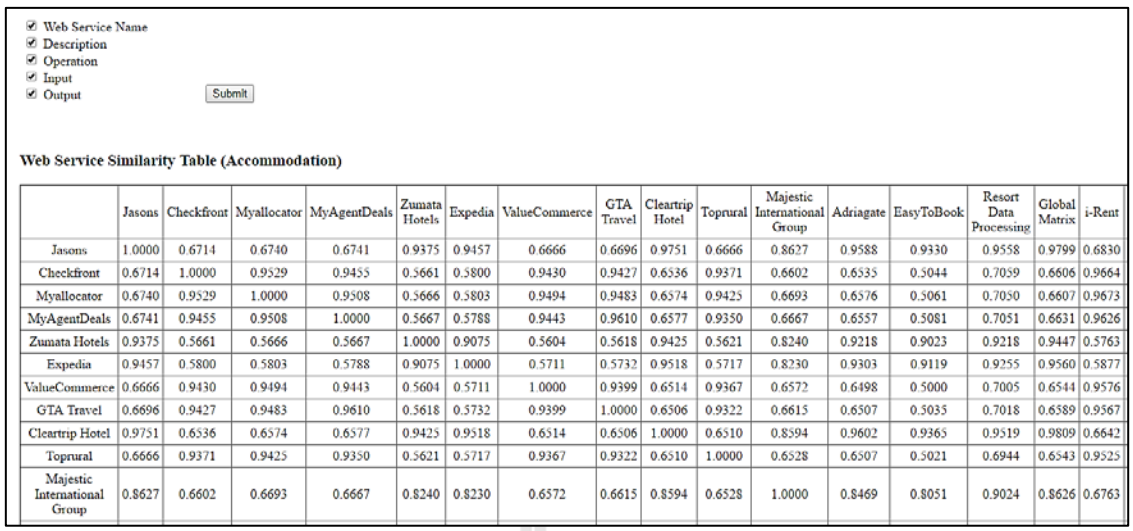
Web Service	Input 1	Input 2	Output 1	Output 2	City	Country	Temp	Wind
A	City	Country	Temp	Wind	1	1	1	1
B	CityName	CouuntryName	Body	-	1	1	0	0
C	City	-	Temp	Wind	1	0	1	1
D	Zip	-	Temp	-	0	0	1	0

หลังจากแทนค่าเวกเตอร์ของคำเสร็จเรียบร้อยแล้วตามตารางที่ 3.7 จากนั้นแทนค่าตามสูตรโคโซน์ตามสมการที่ 3.5 สามารถแสดงตัวอย่างสูตรการคำนวณได้ดังตารางที่ 3.8

ตารางที่ 3.8 การแทนค่าเพื่อคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส รอบที่ 2

Web Service	Similarity Value
A & B	$\text{SIM} = \frac{(1x1)+(1x1)+(1x0)+(1x0)}{\sqrt{(1^2+1^2+1^2+1^2)}\ x\ \sqrt{(1^2+1^2+0^2+0^2)}} = \frac{2}{2x1.41} = \frac{2}{2.82} = 0.71$
A & C	$\text{SIM} = \frac{(1x1)+(1x0)+(1x1)+(1x1)}{\sqrt{(1^2+1^2+1^2+1^2)}\ x\ \sqrt{(1^2+0^2+1^2+1^2)}} = \frac{3}{2x1.73} = \frac{3}{3.46} = 0.87$
A & D	$\text{SIM} = \frac{(1x0)+(1x0)+(1x1)+(1x0)}{\sqrt{(1^2+1^2+1^2+1^2)}\ x\ \sqrt{(0^2+0^2+1^2+0^2)}} = \frac{1}{2x1} = \frac{1}{2} = 0.50$

จากตัวอย่างการคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิสดังแสดงในตารางที่ 3.8 พบว่าเว็บเซอร์วิส C มีค่าความคล้ายคลึงกันกับเว็บเซอร์วิส A มากที่สุด มีค่าเท่ากับ 0.87 รองลงมาคือเว็บเซอร์วิส B มีค่าเท่ากับ 0.71 และเว็บเซอร์วิสที่มีความคล้ายกับเว็บเซอร์วิส A น้อยที่สุดคือเว็บเซอร์วิส D มีค่าเท่ากับ 0.50



☑ Web Service Name
☑ Description
☑ Operation
☑ Input
☑ Output
Submit

Web Service Similarity Table (Accommodation)

	Jasons	Checkfront	Myallocator	MyAgentDeals	Zumata Hotels	Expedia	ValueCommerce	GTA Travel	Cleartrip Hotel	Toprural	Majestic International Group	Adriagate	EasyToBook	Resort Data Processing	Global Matrix	i-Rent
Jasons	1.0000	0.6714	0.6740	0.6741	0.9375	0.9457	0.6666	0.6696	0.9751	0.6666	0.8627	0.9588	0.9330	0.9558	0.9799	0.6830
Checkfront	0.6714	1.0000	0.9529	0.9455	0.5661	0.5800	0.9430	0.9427	0.6536	0.9371	0.6602	0.6535	0.5044	0.7059	0.6606	0.9664
Myallocator	0.6740	0.9529	1.0000	0.9508	0.5666	0.5803	0.9494	0.9483	0.6574	0.9425	0.6693	0.6576	0.5061	0.7050	0.6607	0.9673
MyAgentDeals	0.6741	0.9455	0.9508	1.0000	0.5667	0.5788	0.9443	0.9610	0.6577	0.9350	0.6667	0.6557	0.5081	0.7051	0.6631	0.9626
Zumata Hotels	0.9375	0.5661	0.5666	0.5667	1.0000	0.9075	0.5604	0.5618	0.9425	0.5621	0.8240	0.9218	0.9023	0.9218	0.9447	0.5763
Expedia	0.9457	0.5800	0.5803	0.5788	0.9075	1.0000	0.5711	0.5732	0.9518	0.5717	0.8230	0.9303	0.9119	0.9255	0.9560	0.5877
ValueCommerce	0.6666	0.9430	0.9494	0.9443	0.5604	0.5711	1.0000	0.9399	0.6514	0.9367	0.6572	0.6498	0.5000	0.7005	0.6544	0.9576
GTA Travel	0.6696	0.9427	0.9483	0.9610	0.5618	0.5732	0.9399	1.0000	0.6506	0.9322	0.6615	0.6507	0.5035	0.7018	0.6589	0.9567
Cleartrip Hotel	0.9751	0.6536	0.6574	0.6577	0.9425	0.9518	0.6514	0.6506	1.0000	0.6510	0.8594	0.9602	0.9365	0.9519	0.9809	0.6642
Toprural	0.6666	0.9371	0.9425	0.9350	0.5621	0.5717	0.9367	0.9322	0.6510	1.0000	0.6528	0.6507	0.5021	0.6944	0.6543	0.9525
Majestic International Group	0.8627	0.6602	0.6693	0.6667	0.8240	0.8230	0.6572	0.6615	0.8594	0.6528	1.0000	0.8469	0.8051	0.9024	0.8626	0.6763

รูปที่ 3.13 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Accommodation

ขั้นตอนวิธีการคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส แสดงดังรูปที่ 3.14

```
1   // Generate unique array
2   for ( i = 0; i < count(ws); i++ )
3   {
4           Add each word from ws to unique array
5   }
6
7   // Matrix frequency of words x ws / words x category
8   for ( j = 0; j < count(unique); j++ )
9   {
10          for ( i = 0; i < count(ws); i++ )
11          {
12                  Set word frequency in each ws and object equal 0
13                  Set word frequency in each category and object equal 0
14          }
15  }
16
17  for ( j = 0; j < count(unique); j++ )
18  {
19          for ( i = 0; i < count(ws); i++ )
20          {
21                  Count word frequency in each ws and object
```

```
22                      Count word frequency in each category and object
23              }
24      }
25
26      // Term frequency
27      for ( i = 0; i < count(ws); i++ )
28      {
29              for ( j = 0; j < count(unique); j++ )
30              {
31                      Calculate Term frequency with equation ….
32              }
33      }
34
35      // Inverse document frequency
36      for ( j = 0; j < count(unique); j++ )
37      {
38              Calculate Inverse document frequency with equation ….
39      }
40
41      // TF-IDF
42      for ( i = 0; i < count(ws); i++ )
43      {
44              for ( j = 0; j < count(unique); j++ )
45              {
46                      Calculate inverse document frequency with equation ….
47              }
48      }
49
50      // Calculate cosine similarity
51      for ( i = 0; i < count(ws); i++ )
52      {
53              for ( k = 0; k < count(ws); k++ )
54              {
55                      Calculate weigth with equation ….
56                      Calculate cosine with equation ….
57              }
58      }
```

รูปที่ 3.14 ขั้นตอนวิธีการคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส

3.1.2.5 การหาค่าคุณภาพของเว็บเซอร์วิส (Quality of Service)

วิธีการเลือกเซอร์วิสที่มีความเหมาะสม โดยพิจารณาจาก 5 ปัจจัยดังนี้

1) เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ (Response Time) คือ ระยะเวลาตั้งแต่ผู้ใช้บริการส่งการร้องขอไปยังเซิร์ฟเวอร์ที่ให้บริการจนกระทั่งได้รับข้อมูลตอบกลับมา

2) สภาพความพร้อมในการให้บริการ (Availability) คือ สภาพของเว็บเซอร์วิสที่ปรากฏอยู่และผู้ใช้บริการสามารถเข้าใช้บริการได้ในทันที โดยคำนวณจากเวลาตอบสนองในแต่ละครั้งที่ผู้ใช้บริการสามารถเข้าใช้บริการได้ตามช่วงเวลาที่กำหนด

3) ความเสถียรภาพหรือความน่าเชื่อถือของระบบ (Reliability) คือ ระดับความสามารถในการให้บริการได้ของเว็บเซอร์วิส แสดงได้ด้วยอัตราส่วนของจำนวนครั้งของการเรียกใช้บริการเซอร์วิสที่ประสบความสำเร็จ

4) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) คือ ค่าใช้จ่ายในการเรียกใช้บริการของเซอร์วิส ซึ่งในบางบริการของเว็บเซอร์วิสอาจมีการเก็บค่าบริการ

5) ความสำเร็จในการนำส่งข้อมูล (Throughput) หมายถึง ความสำเร็จในการนำส่งข้อมูลจากจุดหนึ่งไปยังอีกจุดหนึ่งในช่วงเวลาที่กำหนด

วิธีการประเมินเลือกเว็บเซอร์วิสการเลือกเว็บเซอร์วิสที่มีความเหมาะสมของแต่ละการบริการ จะขึ้นอยู่กับคะแนนที่มีค่าสูงสุด รายละเอียดวิธีการวัดค่าทั้ง 5 ปัจจัย และการคำนวณหาค่าคุณภาพการให้บริการของเว็บเซอร์วิสมีดังต่อไปนี้

1) เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ วัดจากระยะเวลาที่ผู้ขอใช้บริการส่งคำร้องขอออกไปจนกระทั่งได้รับข้อมูลตอบกลับมาจากเซิร์ฟเวอร์ของผู้ให้บริการ ซึ่งคำนวณได้จากสมการที่ 3.7

$$T = \frac{\sum_{i=1}^{n} (T_r - T_s)}{n} \qquad (3.7)$$

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

โดยที่ T คือ ค่าเวลาตอบสนองโดยเฉลี่ย หน่วยเป็นมิลลิวินาที

T_r คือ เวลาที่ได้รับข้อมูลกลับมา หน่วยเป็นมิลลิวินาที

T_s คือ เวลาเริ่มต้นที่ทำการส่งคำร้องขอ หน่วยเป็นมิลลิวินาที

RT คือ ค่าความน่าจะเป็นของค่าเวลาตอบสนองโดยเฉลี่ย

ตารางที่ 3.9 ตัวอย่างการคำนวณหาค่าเวลาที่ตอบสนอง

	Web Service 1	Web Service 2	Web Service 3
Round 1	0.4542	0.7189	1.1224
Round 2	0.5798	0.7221	0.9786
Round 3	0.5021	0.5664	0.8519
Round 4	0.3316	0.6039	1.3085
Round 5	0.6122	0.5282	1.7944
Average	0.4960	0.6279	1.2112
RT	**0.7876**	**0.7311**	**0.4813**

2) สภาพความพร้อมในการให้บริการ (Availability) วัดจากอัตราส่วนของผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการสามารถเข้าใช้บริการเซอร์วิสได้สำเร็จส่วนด้วยผลรวมเวลาทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสในช่วงเวลาที่กำหนด ซึ่งคำนวณได้จากสมการที่ 3.8

$$A = \frac{N_{suc}}{N_{all}} \tag{3.8}$$

โดยที่ A คือ ค่าสภาพความพร้อมในการให้บริการ

N_{suc} คือ ผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการสามารถเข้าใช้บริการเซอร์วิสได้สำเร็จ ในช่วงเวลาที่กำหนด มีหน่วยเป็นมิลลิวินาที

N_{all} คือ ผลรวมของเวลาทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิส ในช่วงเวลาที่กำหนด มีหน่วยเป็นมิลลิวินาที

ตารางที่ 3.10 ตัวอย่างการคำนวณหาค่าสภาพความพร้อมในการให้บริการ

	Web Service 1	Web Service 2	Web Service 3
Round 1	**1**	**1**	**1**
Round 2	**7**	**1**	**1**
Round 3	**1**	**1**	**1**
Round 4	**1**	**2**	**1**
Round 5	**1**	**2**	**1**
A	**4/11 = 0.3636**	**3/7 = 0.4286**	**5/5 = 1.0000**

3) ความเสถียรของระบบ (Reliability) วัดจากอัตราส่วนของจำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเซอร์วิสได้สำเร็จส่วนด้วยจำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเซอร์วิส ในช่วงเวลาที่กำหนด ซึ่งคำนวณได้จากสมการที่ 3.9

$$R = \frac{T_{suc}}{T_{all}} \tag{3.9}$$

โดยที่ R คือ ค่าความเสถียรของระบบ

T_{suc} คือ จำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสได้สำเร็จ ในช่วงเวลาที่กำหนด

T_{all} คือ จำนวนครั้งทั้งหมดที่ผู้ใช้บริการเรียกใช้บริการเว็บเซอร์วิสในช่วงเวลาที่กำหนด

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

ตารางที่ 3.11 ตัวอย่างการคำนวณหาค่าความเสถียรของระบบ

	Web Service 1	Web Service 2	Web Service 3
Round 1	Success	Success	Success
Round 2	Fail	Success	Success
Round 3	Success	Success	Success
Round 4	Success	Fail	Success
Round 5	Success	Fail	Success
R	**4/5 = 0.80**	**3/5 = 0.60**	**5/5 = 1.00**

4) ค่าใช้จ่ายในการเรียกใช้บริการ วัดจากอัตราการเก็บค่าบริการในการให้บริการเซอร์วิสของผู้ให้บริการในกลุ่มเดียวกัน ซึ่งคำนวณได้จากสมการที่ 3.10

$$C = 1 - \frac{P}{\sum P_i} \tag{3.10}$$

โดยที่ C คือ ค่าความน่าจะเป็นของค่าใช้จ่ายในการเรียกใช้บริการ

P คือ อัตราการเก็บค่าบริการในการให้บริการเซอร์วิส

$\sum P_i$ คือ ผลรวมของอัตราการเก็บค่าบริการในการให้บริการเซอร์วิสของผู้ให้บริการทั้งหมด

ตารางที่ 3.12 ตัวอย่างการคำนวณหาค่า ค่าใช้จ่ายในการเรียกใช้บริการ

	Web Service 1	Web Service 2	Web Service 3
Cost	10	0 (Free)	5
C	**0.3333**	**1.0000**	**0.6667**

5) Throughput หมายถึง ความสำเร็จในการนำส่งข้อมูลจากจุดหนึ่งไปยังอีกจุดหนึ่งในช่วงเวลาที่กำหนด มีหน่วยเป็นบิตต่อวินาที (bps) Throughput เป็นจำนวนทรานแซคชัน (Transaction) หรือการร้องขอ (Request) ที่ถูกสร้างขึ้นหรือทำงานได้ในช่วงเวลาการทดสอบหนึ่ง ๆ ก่อนเริ่มทำการทดสอบประสิทธิภาพต้องกำหนดช่วงเวลาไว้ก่อนเสมอ ว่าต้องการเท่าไร เช่น จำนวนใน 1 วินาที หรือจำนวนใน 1 นาที เป็นต้น หน่วยที่นิยมใช้สำหรับวัดค่า Throughput คือ TPS (Transaction Per Second) นั่นคือจำนวนทรานแซคชันที่สามารถทำงานได้ใน 1 วินาที ค่าคะแนนคุณภาพของปัจจัยนี้จะบ่งบอกได้ถึงประสิทธิภาพของการให้บริการเว็บเซอร์วิสโดยเฉพาะในเรื่องขนาดของช่องทางที่ใช้ในการสื่อสารข้อมูลหรือแบนด์วิธ (Bandwidth) เพราะการบริการบางอย่างจำเป็นต้องมีแบนด์วิธเพียงพอจึงจะสามารถใช้งานได้อย่างมีประสิทธิภาพ คำนวณได้จากสมการที่ 3.11

$$Tp = \frac{N_{suc}}{T_{Total}} \tag{3.11}$$

$$Tpn = \frac{Tp}{max(Tp)}$$

โดยที่ Tp คือ ค่า Throughput หน่วยเป็น invoke/second

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

T_{total} คือ ค่า ช่วงเวลาการทดสอบ หน่วย second

Tpn คือ ค่า Throughput normalize

$max(Tp)$ คือ ค่า Throughput ที่มีค่าสูงสุด

ตารางที่ 3.13 ตัวอย่างการคำนวณหาค่าความสำเร็จในการนำส่งข้อมูล

	Web Service 1	Web Service 2	Web Service 3
Round 1	**0.15**	**0.21**	**0.23**
Round 2	**0.42**	**0.22**	**0.22**
Round 3	**0.15**	**0.13**	**0.25**
Round 4	**0.25**	**0.64**	**0.21**
Round 5	**0.23**	-	**0.29**
***Tp* (invoke/ 1.2 sec)**	**4/1.2 = 3.33**	**3/1.2 = 2.50**	**5/1.2 = 4.17**
Tpn	**3.33/4.17= 0.80**	**2.50/4.17= 0.60**	**4.17/4.17= 1.00**

6) การคำนวณค่าคุณภาพการให้บริการของเว็บเซอร์วิส การคำนวณค่าคุณภาพการให้บริการของเว็บเซอร์วิสโดยใช้สมการถ่วงน้ำหนัก คำนวณจากการรวมค่าคุณภาพการให้บริการของทั้ง 5 ปัจจัย ได้แก่ เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ สภาพพร้อมใช้งาน ความน่าเชื่อถือ ค่าใช้จ่ายในการเรียกใช้บริการ และความสำเร็จในการนำส่งข้อมูล สูตรการคำนวณดังสมการที่ 3.12

$$QoS_{Si} = \sum_{xeV} W_x \; QoS_{Si,x} \tag{3.12}$$

โดยที่ QoS_{si} คือ ค่าคุณภาพการให้บริการของเว็บเซอร์วิส *Si*

$QoS_{si,x}$ คือ ค่าคุณภาพการให้บริการด้านปัจจัย *x* ของเว็บเซอร์วิส *Si*

W_x คือ ค่าน้ำหนักที่ผู้ใช้บริการกำหนดให้กับคุณภาพด้านปัจจัย *x*

V คือ เซตของปัจจัยที่ใช้กำหนดคุณภาพการให้บริการของเซอร์วิส 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

ตารางที่ 3.14 ค่าคะแนนคุณภาพการให้บริการของเว็บเซอร์วิสหลังถ่วงน้ำหนักแล้ว

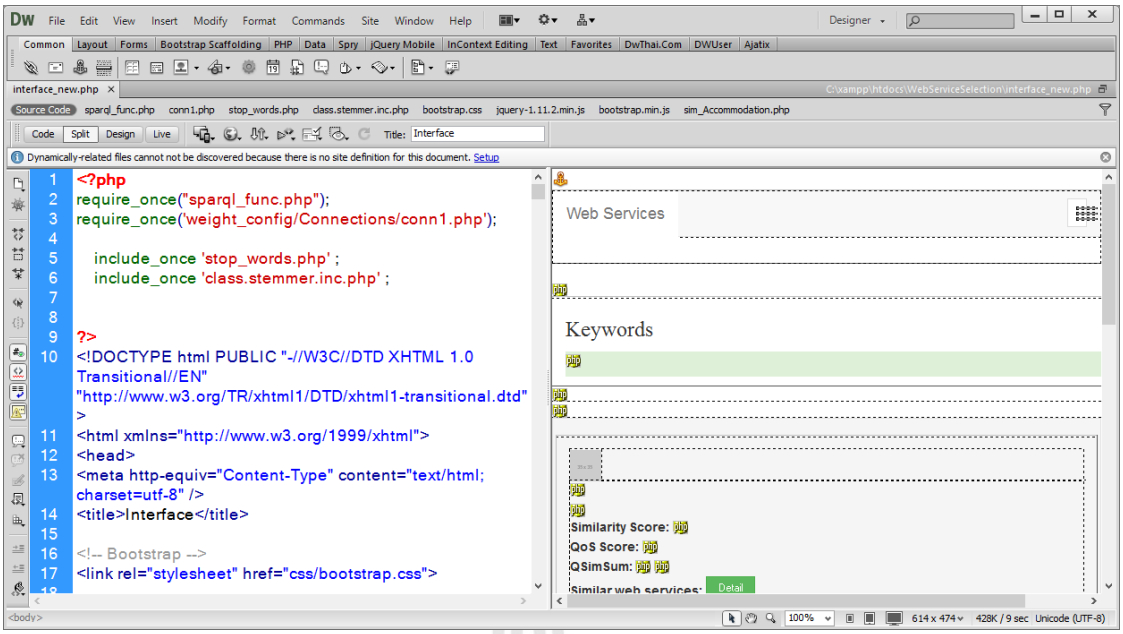
Factors / Web Services	Weight	Web Service 1	Web Service 2	Web Service 3
Response Time	1.00	0.7876	0.7311	0.4813
Availability	0.75	0.3636	0.4286	1.0000
Reliability	0.75	0.8000	0.6000	1.0000
Cost	0.50	0.3333	1.0000	0.6667
Throughput	0.75	0.8000	0.6000	1.0000
QoS		**0.6472**	**0.6540**	**0.8172**

3.1.2.6 การออกแบบและพัฒนากระบวนการค้นหาเว็บเซอร์วิสเชิงความหมาย

กระบวนการในการการออกแบบและพัฒนากระบวนการค้นหาเว็บเซอร์วิสเชิงความหมาย มีขั้นตอนดังต่อไปนี้

1) ออกแบบและพัฒนาส่วนติดต่อกับผู้ใช้ (User Interface) ในการค้นหาเว็บเซอร์วิสเชิงความหมาย

ในการค้นหาข้อมูลเว็บเซอร์วิส ผู้ใช้จะต้องใส่คำสำคัญ (Keyword) เพื่อสืบค้นเข้าไปในระบบ จากนั้นระบบจะทำการเปรียบเทียบระหว่างคำสำคัญกับฐานความรู้ออนโทโลยีของเว็บเซอร์วิส ดังนั้นจึงต้องมีการพัฒนาส่วนติดต่อกับผู้ใช้ เพื่อเป็นตัวกลางในการติดต่อระหว่างผู้ใช้และฐานความรู้ออนโทโลยี โดยใช้ภาษาพีเอชพี (PHP) เป็นภาษาหลักในการพัฒนาส่วนติดต่อกับผู้ใช้ ใช้ซีเอสเอส (Cascading Style Sheet : CSS) ในการจัดรูปแบบการแสดงผล ส่วนการสืบค้นข้อมูลจากออนโทโลยีใช้ภาษาสปาร์เคิล (SPARQL) เพื่อสอบถามข้อมูล (Query) และดึงข้อมูลจากโอดับเบิ้ลยูแอล (OWL) ผ่านทางแรพเอพีไอ (RAP API)



รูปที่ 3.15 โปรแกรมอะโดบีดรีมเวฟเวอร์สำหรับออกแบบส่วนติดต่อกับผู้ใช้

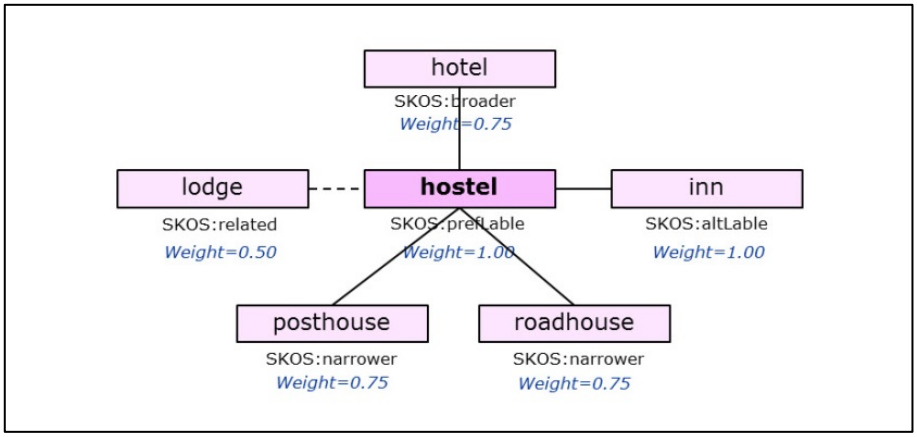
2) การกำหนดค่าน้ำหนักของคำในโครงสร้างสคอส (SKOS)

ในการกำหนดค่าน้ำหนักที่ถูกจัดเก็บตามโครงสร้างแบบสคอส (SKOS) ซึ่งค่าน้ำหนักนี้เป็นการกำหนดค่าเริ่มต้น ทั้งนี้ผู้ใช้สามารถกำหนดเองได้ภายหลัง ซึ่งสามารถกำหนดได้ในเมนู “Setting” บนหน้าเว็บหลัก การกำหนดค่าน้ำหนักเริ่มต้นแสดงได้ดังตารางที่ 3.15

ตารางที่ 3.15 ตัวอย่างการให้ค่าน้ำหนักของคำตามโครงสร้างสคอส (SKOS)

ประเภทของคำ	SKOS Label	ค่าน้ำหนัก
คำหลัก	prefLabel	1.00
คำเหมือน	altLabel	1.00
คำที่กว้างกว่า	broader	0.75
คำที่แคบกว่า	narrower	0.75
คำที่เกี่ยวข้อง	related	0.50

ตัวอย่างการนิยามความสัมพันธ์ของคำพร้อมค่าน้ำหนัก ในตัวอย่างนี้มีคำหลักคือ“hostel” คำเหมือนคือ “inn” คำที่กว้างกว่าคือ “hotel” คำที่แคบกว่าคือ “posthouse” และ “roadhouse” และสุดท้ายคำที่เกี่ยวข้องคือ “lodge” แสดงดังรูปที่ 3.16



รูปที่ 3.16 ตัวอย่างโครงสร้างความสัมพันธ์ในรูปแบบสคอสของคำหลัก “hostel” (SKOS Vocabulary Relationship)

3) การกำหนดค่าน้ำหนักของคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส

ในการกำหนดค่าน้ำหนักของคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส ซึ่งคำเหล่านี้ถูกจัดเก็บไว้ในรูปแบบของคลาสในออนโทโลยีแล้ว ได้แก่ ชื่อเว็บเซอร์วิส (Web Service Name) คำอธิบายเว็บเซอร์วิส (Description) ชื่อโอเปอเรชัน (Operation) ข้อมูลเข้า (Input) และข้อมูลออก (Output) ทั้งนี้ผู้ใช้สามารถกำหนดค่าน้ำหนักได้เองตามความเหมาะสมในภายหลังก่อนที่จะค้นหาเว็บเซอร์วิส ซึ่งสามารถกำหนดได้ในเมนู “Setting” บนหน้าเว็บหลัก การกำหนดค่าน้ำหนักเริ่มต้นแสดงได้ดังตารางที่ 3.16

ตารางที่ 3.16 ตัวอย่างการกำหนดค่าน้ำหนักของคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส

ชื่อตำแหน่ง	ค่าน้ำหนัก
Web service name	0.50
Description	1.00
Operation	0.50
Input	0.50
Output	0.75

4) การกำหนดค่าน้ำหนักของคำที่ใช้สำหรับสืบค้น

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

ตัวอย่างการคำนวณหาค่าความคล้ายคลึงระหว่างคำสำคัญกับออนโทโลยีเว็บเซอร์วิส มีขั้นตอนดังนี้

ก) กำหนดให้คำสำคัญคือ “hostel booking” ดังนั้นค่าน้ำหนักตามตำแหน่งของคำสำคัญนี้คือ คำสำคัญในลำดับแรกคำว่า “hostel” มีค่าน้ำหนักคือ 1.00 และคำสำคัญในลำดับที่สองคำว่า “booking” มีค่าน้ำหนักคือ 0.75

ข) นำกลุ่มของคำสำคัญที่ได้กำหนดค่าน้ำหนักแล้วไปเปรียบเทียบกับ index ในคลังคำที่ถูกจัดเก็บตามโครงสร้างแบบ SKOS จากนั้นจะได้กลุ่มคำพร้อมค่าน้ำหนักดังนี้

hostel = { hostel(1.00), inn(1.00), hotel(0.75), posthouse(0.75),
roadhouse(0.75), lodge(0.50) }

booking = { booking(1.00), reservation(1.00), employment(0.75),
reserve(0.50), book(0.50) }

ค) คำนวณหาค่า TF-IDF ให้กับคำทุกคำในออนโทโลยีเว็บเซอร์วิสตามตำแหน่งโครงสร้างของเว็บเซอร์วิสตามสมการที่ 3.13 และสมการที่ 3.14

$$tf(t,d) = 0.5 + \frac{0.5 * f(t,d)}{\max\{f(w,d) : w \in d\}} \quad (3.13)$$

โดยที่ $tf(t,d)$ คือ ค่าน้ำหนักของคำ t บนเว็บเซอร์วิส d

$f(t,d)$ คือ จำนวนความถี่ของคำ t บนเว็บเซอร์วิส d

$\max\{f(w,d) : w \in d\}$ คือ ความถี่สูงสุดของคำ w บนเว็บเซอร์วิส d

$$idf(t,D) = log \frac{|D|}{|\{d \in D : t \in d\}|} \quad (3.14)$$

โดยที่ $idf(t,d)$ คือ ส่วนกลับค่าน้ำหนักของคำ t บนเว็บเซอร์วิสทั้งหมด

$|D|$ คือ จำนวนของเว็บเซอร์วิสทั้งหมด

$|\{d \in D : t \in d\}|$ คือ จำนวนเว็บเซอร์วิสที่มีคำ t

$$WSP = \sum_{k=0}^{n} (tf_idf_k * WePoWS_k) \quad (3.15)$$

โดยที่ *WSP* คือ ค่าน้ำหนักของคำที่อยู่ในแต่ละตำแหน่งตามโครงสร้างของเว็บเซอร์วิส

tf_idf คือ ค่าน้ำหนักความสำคัญของคำ

WePoWS (Weight of Position Web Service) คือ ค่าน้ำหนักประจำตำแหน่งในโครงสร้างของเว็บเซอร์วิส ค่าน้ำหนักแสดงดังตารางที่ 3.16

5) คำนวณหาค่าน้ำหนักของคำทุกคำที่อยู่ในกลุ่มคำที่ได้จากข้อ 3) เมื่อเปรียบเทียบกับค่าน้ำหนักที่ถูกจัดเก็บตามโครงสร้างแบบสคอส (SKOS) ค่าน้ำหนักแสดงดังตารางที่ 3.15 สามารถแทนค่าได้ตามสมการที่ 3.16

$$SKOSPOS = \sum_{k=0}^{n} (WSP_k * WePoSK_k) \quad (3.16)$$

โดยที่ *SKOSPOS* คือ ค่าน้ำหนักของคำที่อยู่ในโครงสร้างของสคอส (SKOS)

WSP คือ ค่าน้ำหนักของคำที่อยู่ในแต่ละตำแหน่งตามโครงสร้างของเว็บเซอร์วิส

WePoSK (Weight of Position SKOS) คือ ค่าน้ำหนักประจำตำแหน่งในโครงสร้างของเว็บเซอร์วิส

6) การคำนวณหาค่าความคล้ายคลึงกันเชิงความหมายระหว่างคำสำคัญกับออนโทโลยีเว็บเซอร์วิส สามารถแทนค่าได้ตามสมการดังนี้

$$SIM = \sum_{k=0}^{n} (SKOSPOS_k * WPQ_k) \quad (3.17)$$

โดยที่ *SIM* คือ ค่าคะแนนเชิงความหมายของแต่ละเว็บเซอร์วิส

SKOSPOS คือ ค่าน้ำหนักของคำที่อยู่ในโครงสร้างของ SKOS

WPQ (Weight of Position Query) คือ ค่าน้ำหนักประจำตำแหน่งของคำสำคัญที่ได้จากผู้ใช้

$$SEMSIM = \frac{SIM_k}{\max(SIM)} \quad (3.18)$$

โดยที่ *SEMSIM* คือ ค่าคะแนนความคล้ายคลึงกันเชิงความหมายหลังจากผ่านการปรับปรุงค่า (Normalization) โดยให้ค่าคะแนนอยู่ระหว่าง 0 ถึง 1

SIM คือ ค่าคะแนนเชิงความหมายของแต่ละเว็บเซอร์วิส

max(*SIM*) คือ ค่าสูงสุดของค่าคะแนนเชิงความหมายของเว็บเซอร์วิสทั้งหมด

7) ขั้นตอนการค้นหาเว็บเซอร์วิสเชิงความหมาย ผู้วิจัยทำการออกแบบขั้นตอนการค้นหาเว็บเซอร์วิสเชิงความหมาย มีขั้นตอนดังต่อไปนี้

ก) รับคำค้น (Keyword) จากผู้ใช้ แล้วมาแยกคำค้นด้วยช่องว่าง

ข) กำหนดค่าน้ำหนักให้กับ QoS น้ำหนักคำตามลำดับคำค้น น้ำหนักคำตามโครงสร้างสคอส และน้ำหนักคำตามตำแหน่งโครงสร้างของเว็บเซอร์วิส

ค) นำคำสำคัญ (Keyword) ไปหากลุ่มคำด้วย Stem จากรายการคำใน JSON

ง) นำ กลุ่มคำ จากข้อ ค) ที่แยกไว้แต่ละคำ ไปหา prefLabel, altLabel, broader, narrower และ related จากไฟล์ OWL

จ) อ่านค่า TF-IDF ของคำทั้งหมดที่อยู่ JSON (JSON อ่านจาก SKOS) ที่คำนวณไว้แล้ว

ฉ) เอาผลลัพธ์จากข้อ ค) มาเปรียบเทียบหาค่า TF-IDF จากข้อ ง)

ช) นำผลลัพธ์จากข้อ จ) มาถ่วงน้ำหนักตามตำแหน่งโครงสร้างของเว็บเซอร์วิส

ซ) นำผลลัพธ์จากข้อ ฉ) มาถ่วงน้ำหนักตามโครงสร้างสคอส

ฌ) นำผลลัพธ์จากข้อ ช) มาถ่วงน้ำหนักตามลำดับคำค้น

ญ) เรียงลำดับค่าความคล้ายจากข้อ ซ) จากมากไปน้อย และลบค่าความคล้ายที่เป็น 0 หรือน้อยที่สุดออก เพราะค่า 0 คือการค้นหาแล้วไม่เจอผลลัพธ์

ฎ) อ่านค่า QoS ที่คำนวณไว้แล้ว

ฏ) นำผลลัพธ์จากข้อ ญ) มาถ่วงน้ำหนักค่าคะแนนคุณภาพ (QoS) ในทุกปัจจัย

ฐ) แสดงผลโดยเรียงลำดับจากค่า Sim หรือ QoS หรือ Sim+QoS มากไปน้อย ตามสมการที่ 3.19 หรือ Ranksum ตามสมการที่ 3.20

มหาวิทยาลัยเทคโนโลยีสุรนารี

$$QSS = \left[\frac{(Q*W_Q)+(S*W_S)}{W_Q+W_S}\right] \quad (3.19)$$

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)

ตารางที่ 3.17 ตัวอย่างค่าคะแนน Similarity และค่าคะแนน QoS

Web Service	Similarity Score	QoS Score
A	0.75	0.66
B	0.98	0.87
C	0.54	0.91

ตารางที่ 3.18 ค่าน้ำหนัก Similarity และค่าน้ำหนัก QoS

Similarity Weight	Qos Weight
1.0	1.0

ตารางที่ 3.19 ผลลัพธ์ของค่าคะแนน QsimSum ของแต่ละเว็บเซอร์วิส

Web Service	QSimSum
A	0.705
B	0.925
C	0.725

$$RankSum(W_i) = \frac{n-r_j+1}{\sum(n-r_k+1)} \quad (3.20)$$

โดยที่ Wi คือ ค่าน้ำหนักของข้อมูล

r_j คือ ลำดับความสำคัญของข้อมูล

r_k คือ ผลรวมของค่าของข้อมูล

n คือ จำนวนข้อมูลทั้งหมด

ตารางที่ 3.20 ตัวอย่างค่าคะแนน Similarity และค่าคะแนน QoS

Web Service	Similarity Score	QoS Score
A	0.75	0.66
B	0.98	0.87
C	0.54	0.91

ตารางที่ 3.21 ค่าลำดับของแต่ละเว็บเซอร์วิส

Web Service	Similarity Rank	QoS Rank
A	2	3
B	1	2
C	3	1

ตารางที่ 3.22 ค่าส่วนกลับลำดับของ Similarity และ QoS

Web Service	Similarity Inverse Rank	QoS Inverse Rank
A	2	1
B	3	2
C	1	3

ตารางที่ 3.23 ค่าคะแนนของ Similarity และ QoS

Web Service	Similarity Inverse Rank	QoS Inverse Rank
A	0.3333	0.1667
B	0.5000	0.3333
C	0.1667	0.5000

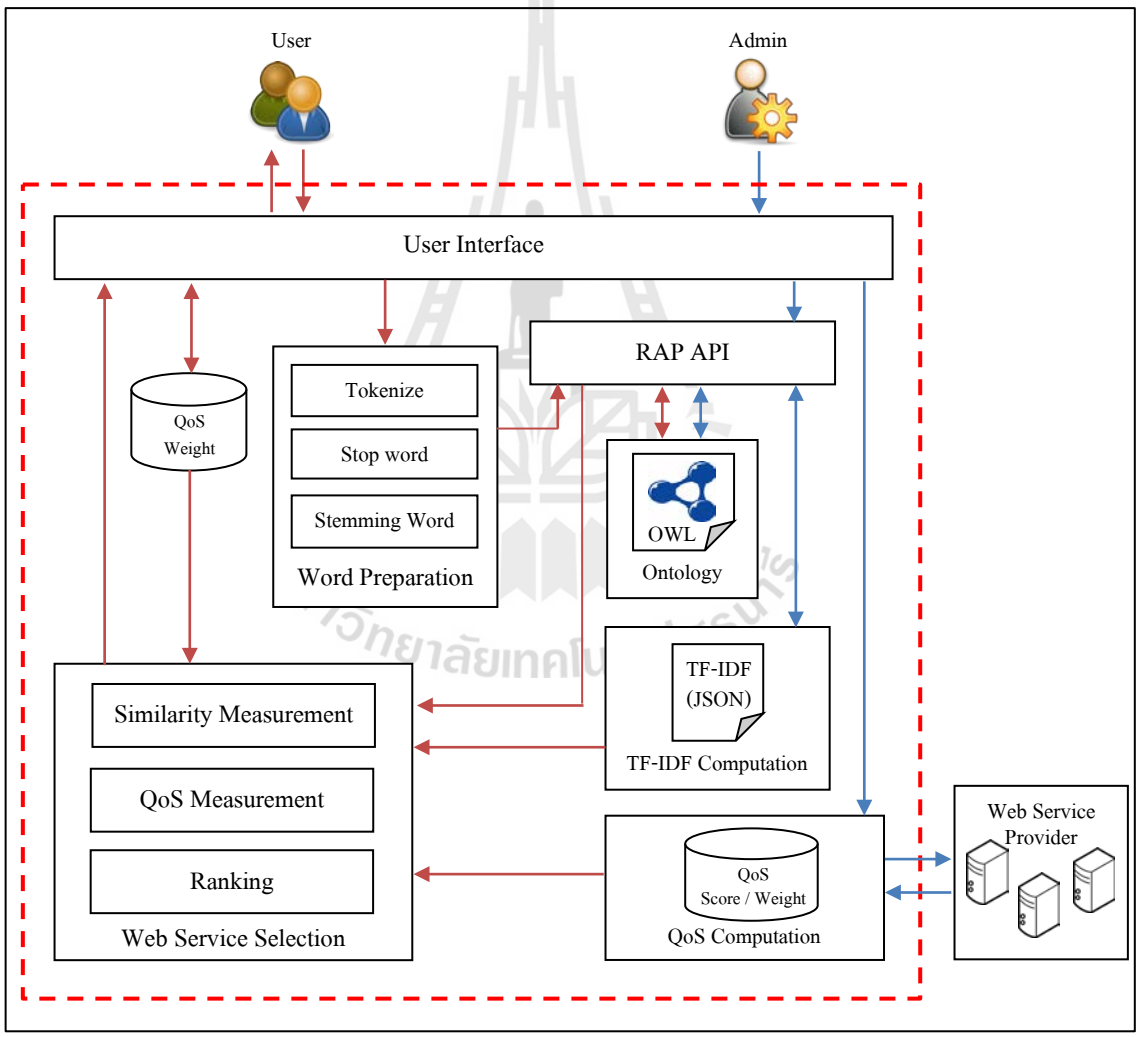
ตารางที่ 3.24 ค่าคะแนน Ranksum ก่อนถ่วงน้ำหนัก

Web Service	Ranksum before weight
A	0.0556
B	0.1667
C	0.0833

ตารางที่ 3.25 ค่าคะแนน Ranksum

Web Service	Ranksum
A	0.3333
B	1.0000
C	0.5000

8) ขั้นตอนการค้นหาเว็บเซอร์วิสเชิงความหมาย สามารถสรุปออกมาเป็นสถาปัตยกรรมของระบบค้นหาเว็บเซอร์วิสเชิงความหมาย แสดงดังรูปที่ 3.17



รูปที่ 3.17 สถาปัตยกรรมของระบบค้นหาเว็บเซอร์วิสเชิงความหมาย

3.1.3 การประเมินผล

3.1.3.1 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 และไม่เกี่ยวข้องกับสิ่งที่ต้องการ (True Negative: *TN*)

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

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

ตารางที่ 3.26 คอนฟิวส์ชันเมทริกซ์ (Confusion Matrix) ของผลการสืบค้น

ประเภท	ผลการสืบค้น	
	สืบค้นได้	สืบค้นไม่ได้
เอกสารที่เกี่ยวข้อง	*TP*	*FN*
เอกสารที่ไม่เกี่ยวข้อง	*FP*	*TN*

จากการสืบค้นเว็บเซอร์วิสจะเห็นได้ว่าการสืบค้นนั้นไม่สามารถค้นคืนเว็บเซอร์วิสที่เกี่ยวข้องกับสิ่งที่ผู้ใช้ต้องการได้ 100% จึงมีการคิดค่าการประเมินผลการสืบค้นเว็บเซอร์วิสได้แก่ ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าเอฟเมเซอร์ (F-Measure) (Miao, Duan, Zhang, and Jiao, 2009: p.9172)

ค่าความแม่นยำ (Precision) คือ การพิจารณาความถูกต้องของข้อมูลที่สืบค้นได้ จากอัตราส่วนระหว่างจำนวนเว็บเซอร์วิสที่ค้นคืนถูกต้องกับจำนวนเว็บเซอร์วิสที่ค้นคืนมาได้ทั้งหมด

(ตัวอย่าง: เมื่อค่าความแม่นยำเท่ากับ 1 แสดงว่าเว็บเซอร์วิสที่ค้นคืนมาได้ทั้งหมดเป็นผลลัพธ์ที่ถูกต้อง) สูตรการหาค่าความแม่นยำแสดงดังสมการที่ 3.21

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.21}$$

โดยที่ Precision คือ ค่าความแม่นยำในการสืบค้นเว็บเซอร์วิส
TP คือ เว็บเซอร์วิสที่เกี่ยวข้องและถูกสืบค้นได้
FP คือ เว็บเซอร์วิสที่ไม่เกี่ยวข้องแต่สืบค้นได้

ค่าความระลึก (Recall) คือ การพิจารณาความครบถ้วนของข้อมูลเมื่อเทียบกับข้อมูลที่ควรได้ทั้งหมด จากอัตราส่วนระหว่างจำนวนเว็บเซอร์วิสที่ค้นคืนถูกต้องกับจำนวนเว็บเซอร์วิสที่ถูกต้องทั้งหมดในระบบ (หมายเหตุ: ค่าความระลึกเน้นให้คำตอบที่ถูกต้องออกมาให้ได้มากที่สุด) สูตรการหาค่าความระลึกแสดงดังสมการที่ 3.22

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.22}$$

โดยที่ Recall คือ ค่าความระลึกในการสืบค้นเว็บเซอร์วิส
TP คือ เว็บเซอร์วิสที่เกี่ยวข้องและถูกสืบค้นได้
FN คือ เว็บเซอร์วิสที่เกี่ยวข้องแต่สืบค้นไม่ได้

นอกจากการทดสอบประสิทธิภาพของระบบเพื่อวัดความถูกต้องและแม่นยำด้วยค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) แล้ว ยังมีการวัดคุณภาพของผลการค้นหาข้อมูลโดยรวมด้วยค่าเอฟเมเชอร์ (F-Measure) ซึ่งเป็นค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก แสดงได้ดังสมการที่ 3.23

$$\text{F-measure} = 2\ \frac{\text{precison} * \text{recall}}{\text{precison} + \text{recall}} \tag{3.23}$$

โดยที่ F-measure คือ ค่าเอฟเมเชอร์ที่วัดคุณภาพของผลการค้นหาโดยรวม
precision คือ ค่าความแม่นยำในการสืบค้นเว็บเซอร์วิส
recall คือ ค่าความระลึกในการสืบค้นเว็บเซอร์วิส

3.1.3.2 การประเมินผลการจำแนกหมวดหมู่ของเว็บเซอร์วิส

โดยใช้การทำเหมืองข้อมูล (Data Mining) และใช้เทคนิควิธีทั้ง 5 เทคนิคนี้ ได้แก่

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

ข) เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) โดยกำหนดค่า cross-validation เท่ากับ 10

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

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

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

จากนั้นประเมินผลการจำแนกหมวดหมู่ของเว็บเซอร์วิส โดยการคำนวณหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ตามสมการที่ 3.21 สมการที่ 3.22 สมการที่ 3.23 และตามสมการที่ 3.24

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{3.24}$$

โดยที่ Accuracy คือ ค่าความถูกต้องในการสืบค้นเว็บเซอร์วิส

TP คือ เว็บเซอร์วิสที่เกี่ยวข้องและถูกสืบค้นได้

TN คือ เว็บเซอร์วิสที่ไม่เกี่ยวข้องและสืบค้นไม่ได้

FP คือ เว็บเซอร์วิสที่ไม่เกี่ยวข้องแต่สืบค้นได้

FN คือ เว็บเซอร์วิสที่เกี่ยวข้องแต่สืบค้นไม่ได้

3.1.3.3 การประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการ

ก่อนการประเมินผลผู้วิจัยได้ทำการหาจำนวนคำที่เหมาะสมในการประเมินประสิทธิภาพในการค้นหาเว็บเซอร์วิสเชิงความหมายโดยใช้ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation: SD) เพื่อหาจำนวนคำที่เหมาะสม โดยการสุ่มคำค้นในฐานข้อมูลมาจำนวน 5 ชุดตัวอย่าง แล้วนำคำค้นที่สุ่มได้ไปค้นหาเว็บเซอร์วิส จากนั้นนำค่าเอฟเมเชอร์ (F-measure) ที่ได้จาก 5 ชุดตัวอย่างมาหาค่าเบี่ยงเบนมาตรฐาน จากนั้นจะได้จำนวนคำที่เหมาะสมเพื่อใช้ในการประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการต่อไปดังขั้นตอนดังต่อไป

1) ประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการโดยการเปรียบเทียบผลลัพธ์ที่ได้ระหว่างผู้เชี่ยวชาญและระบบสืบค้น

ก) กำหนดเว็บเซอร์วิสและคำค้นให้ผู้เชี่ยวชาญระบุว่า ถ้าใช้คำค้นนั้น ๆ แล้วจะได้ผลลัพธ์เป็นเว็บเซอร์วิสใดบ้าง

ข) ทำการสืบค้นด้วยคำค้นเดียวกันจากระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายที่พัฒนาขึ้น นำผลลัพธ์ที่ได้มาคำนวณค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) โดยการจัดผลลัพธ์การสืบค้นให้อยู่ในรูป True Table ซึ่งประกอบด้วย *TP* = True Positive, *TN* = True Negative, *FP* = False Positive และ *FN* = False Negative ตามสมการที่ 3.21 และสมการที่ 3.22

ค) จากนั้นนำค่าผลลัพธ์ที่ได้ มาคำนวณหาค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ตามสมการที่ 3.23 และสมการที่ 3.24

2) ประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการโดยใช้เกณฑ์อัตโนมัติ

ก) ทำการกำหนดคำค้น เพื่อสืบค้นแบบเปรียบเทียบคำสำคัญ (Keyword Matching) เช่น ค้นด้วยคำ A ได้ผลลัพธ์ wsA, wsB, wsD และ wsE เป็นต้น

ข) นำคำค้นเดียวกันมาสืบค้นในระบบที่พัฒนาขึ้น เช่น ค้นด้วยคำ A ได้ผลลัพธ์ wsB 1.0, wsD 0.9, wsC 0.8, wsE 0.7, wsA 0.6 และ wsF 0.5 เป็นต้น

ค) นำผลลัพธ์จากการสืบค้นในข้อ ก) มาเป็นข้อกำหนดในการตัดเกณฑ์ในผลลัพธ์การสืบค้นของข้อ ข) ว่าจะเลือกตัดที่ค่าต่ำสุดที่ถูกค้นพบในการสืบค้นในข้อ ก) เช่น ข้อ ก) พบ wsA, wsB, wsD และ wsE แล้วได้คะแนนในข้อ ข) 0.6, 1.0, 0.9 และ 0.7 ตามลำดับ เลือกค่าต่ำสุดเป็นเกณฑ์คือ 0.6

ง) นำผลลัพธ์ที่ได้มาจัดผลลัพธ์การสืบค้นให้อยู่ในรูป True Table ซึ่งประกอบ *TP* = True Positive, *TN* = True Negative, *FP* = False Positive และ *FN* = False Negative

จ) คำนวณหาค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) และค่าเอฟเมเชอร์ (F-measure) ตามสมการที่ 3.21 สมการที่ 3.22 และสมการที่ 3.23

3.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวมีรายละเอียดดังนี้

1) เครื่องคอมพิวเตอร์ส่วนบุคคลชนิดพกพาจำนวน 1 เครื่อง โดยมีคุณสมบัติดังนี้
   - หน่วยประมวลผลกลางชนิด Intel Core i7 ความเร็ว 3.60 GHz
   - หน่วยความจำหลักขนาด 8 GB
   - หน่วยความจำสำรองขนาด 1 TB
   - เครื่องพิมพ์ชนิดเลเซอร์ จำนวน 1 เครื่อง
2) เครือข่ายคอมพิวเตอร์และช่องสัญญาณอินเทอร์เน็ตความเร็ว DL.100 / UL.50 Mbps
3) ระบบปฏิบัติการและโปรแกรมประยุกต์
   - ระบบปฏิบัติการ Windows 10 Pro
   - เว็บบราวเซอร์ Google Chrome, Mozilla Firefox และ Safari
   - โปรแกรมโพรเทเจ (Protégé)
   - โปรแกรม Adobe Dreamweaver
   - โปรแกรม EditPlus
   - โปรแกรมภาษา PHP
   - ฐานข้อมูล MySQL
   - โปรแกรมจัดการฐานข้อมูล PhpMyAdmin
   - โปรแกรม WEKA
   - โปรแกรม RapidMiner Studio

3.3 การวิเคราะห์ข้อมูล

3.3.1 การวิเคราะห์ข้อมูลสำหรับการจำแนกหมวดหมู่ของเว็บเซอร์วิสใช้วิธีการทำเหมืองข้อมูล (Data Mining) ซึ่งใช้เทคนิคต่าง ๆ ได้แก่ เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) เทคนิควิธีเค-เนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) และเทคนิควิธีเคมีนส์ (K-means) จากนั้นประเมินผลการจำแนกหมวดหมู่ของเว็บเซอร์วิสโดยใช้สูตรการหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าเอฟเมเซอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ตามสมการที่ 3.21 สมการที่ 3.22 สมการที่ 3.23 และสมการที่ 3.24 ตามลำดับ

3.3.2 การวิเคราะห์ข้อมูลเพื่อการคำนวณหาค่าความคล้ายคลึงกันของเว็บเซอร์วิส (Web Services Similarity) ขั้นตอนนี้จะทำการคำนวณเพื่อหาค่าความคล้ายคลึงกันของเว็บเซอร์วิสในแต่ละกลุ่ม โดยใช้วิธีการหาค่าความคล้ายคลึงตามทฤษฎี Vector Space Model (VSM) คือการวัดความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) โดยพิจารณาเปรียบเทียบความคล้ายคลึงกันจากชื่อของเว็บเซอร์วิส (Web Service Name) ชื่อโอเปอเรชัน (Operation Name) คำอธิบาย (Description) ข้อมูลเข้า (Input) และข้อมูลออก (Output) ตามสมการที่ 3.5

3.3.3 การวิเคราะห์ข้อมูลเพื่อการคำนวณหาค่าคุณภาพของเว็บเซอร์วิสทั้ง 5 ปัจจัย ได้แก่ เวลาที่ตอบสนองของเซิร์ฟเวอร์ที่ให้บริการ (Response Time) สภาพความพร้อมในการให้บริการ (Availability) ความเสถียรภาพหรือความน่าเชื่อถือ (Reliability) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) และความสำเร็จในการนำส่งข้อมูล (Throughput) ตามสมการที่ 3.7 สมการที่ 3.8 สมการที่ 3.9 สมการที่ 3.10 และสมการที่ 3.11

3.3.4 การวิเคราะห์ข้อมูลสำหรับการประเมินผลการสืบค้นเว็บเซอร์วิสเชิงความหมาย เมื่อทำการสืบค้นเว็บเซอร์วิสจากออนโทโลยี ระบบจะแบ่งกลุ่มของเว็บเซอร์วิสออกเป็น 2 ส่วน คือ เว็บเซอร์วิสที่ถูกดึงออกมาเป็นผลลัพธ์จากการสืบค้น (Retrieved) และเว็บเซอร์วิสที่ไม่ถูกดึงออกมา (Not Retrieved) ซึ่งเว็บเซอร์วิสต่าง ๆ ใน 2 กลุ่มนี้ อาจมีทั้งเว็บเซอร์วิสที่เกี่ยวข้องมีคุณสมบัติผ่านเกณฑ์การเปรียบเทียบความใกล้เคียงของการสืบค้นครั้งนั้น คือเว็บเซอร์วิสที่ผู้ใช้ต้องการ (Relevant) และไม่เกี่ยวข้อง (Irrelevant) กับสิ่งที่ต้องการ โดยใช้สูตรการหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ตามสมการที่ 3.21 สมการที่ 3.22 สมการที่ 3.23 และสมการที่ 3.24

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 4
ผลการวิจัยและอภิปรายผล

ในบทนี้กล่าวถึงผลการวิจัยและอภิปรายผลของแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว ประกอบไปด้วย ผลการจำแนกหมวดหมู่ของเว็บเซอร์วิสและการประเมินประสิทธิภาพ และผลการคัดเลือกเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยวและการประเมินประสิทธิภาพ ซึ่งมีรายละเอียดดังนี้

4.1 ผลการจำแนกหมวดหมู่ของเว็บเซอร์วิสและการประเมินประสิทธิภาพ

ในการจำแนกหมวดหมู่ของเว็บเซอร์วิสใช้หลักการทำเหมืองข้อมูล (Data Mining)โดยใช้เทคนิควิธีในการทดลองจำนวน 5 เทคนิควิธี ดังนี้

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

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

3) เทคนิควิธีเคเนียร์เรสท์เนเบอร์ (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

หมายเหตุ: เทคนิควิธีที่ 2 - 5 ได้ทำการทดลองเพื่อหาค่า cross-validation ที่เหมาะสม โดยกำหนดค่า cross-validation ระหว่าง 3 ถึง 20 พบว่าค่า cross-validation เท่ากับ 10 มีความเหมาะสมที่สุด เนื่องจากให้ค่าความถูกต้องในการจำแนกหมวดหมู่ของเว็บเซอร์วิสเฉลี่ยสูงที่สุด

การจำแนกหมวดหมู่ของเว็บเซอร์วิส เพื่อเป็นการแสดงให้เห็นถึงประสิทธิภาพของการใช้ฐานความรู้แบบออนโทโลยีในการจัดการข้อมูลเว็บเซอร์วิสเชิงความหมาย โดยผู้วิจัยได้ทำการเก็บรวบรวมเว็บเซอร์วิสจากจำนวนทั้งหมด 200 เว็บเซอร์วิส ทำการคัดเลือกเฉพาะเว็บเซอร์วิสที่มีความสมบูรณ์ที่สุดจำนวน 76 เว็บเซอร์วิส เพื่อนำมาใช้ในการจำแนกหมวดหมู่ของเว็บเซอร์วิส จากนั้นนำ

ข้อมูลเว็บเซอร์วิสที่ได้ผ่านกระบวนการต่าง ๆ เช่น การตัดแบ่งคำ (Word Segmentation) การตัดคำหยุด (Stop words) และการหารากศัพท์ (Stemming Words) เป็นต้น โดยแบ่งข้อมูลของเว็บเซอร์วิสออกเป็น 2 รูปแบบ คือ ข้อมูลในชุดที่ 1 คือข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย และข้อมูลในชุดที่ 2 คือ ข้อมูลของเว็บเซอร์วิสที่ผ่านการหาความสัมพันธ์ของคำในออนโทโลยี จากนั้นทำการเปรียบเทียบการจำแนกข้อมูลเว็บเซอร์วิสระหว่างข้อมูลทั้งสองชุด ผลการทดลองสามารถแสดงดังนี้

ข้อมูลชุดที่ 1: ข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำ

ตารางที่ 4.1 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคมีนส์ (K-means)

Accuracy: 64.47% (K-means)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	17	2	2	1	0	77.27%
pred. Transportation	0	6	2	4	0	50.00%
pred. Attraction	6	3	3	3	0	20.00%
pred. Location	0	3	0	5	0	62.50%
pred. Weather	0	0	0	1	18	94.74%
Recall	73.91%	42.86%	42.86%	35.71%	100.00%	

ตารางที่ 4.2 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB)

Accuracy: 69.46% (Naive Bayes : NB)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	20	3	3	5	2	60.61%
pred. Transportation	3	11	1	2	0	64.71%
pred. Attraction	0	0	1	0	0	100.00%
pred. Location	0	0	2	6	1	66.67%
pred. Weather	0	0	0	1	15	93.75%
Recall	86.96%	78.57%	14.29%	42.86%	83.33%	

ตารางที่ 4.3 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN)

Accuracy: 81.61% (k-Nearest Neighbor : k-NN)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	23	3	2	2	0	76.67%
pred. Transportation	0	11	0	2	0	84.62%
pred. Attraction	0	0	3	1	0	75.00%
pred. Location	0	0	1	7	0	87.50%
pred. Weather	0	0	1	2	18	85.71%
Recall	100.00%	78.57%	42.86%	50.00%	100.00%	

ตารางที่ 4.4 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT)

Accuracy: 82.86% (Decision Tree : DT)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	22	1	1	0	0	91.67%
pred. Transportation	0	9	0	1	0	90.00%
pred. Attraction	1	0	5	2	0	62.50%
pred. Location	0	4	1	11	2	61.11%
pred. Weather	0	0	0	0	16	100.00%
Recall	95.65%	64.29%	71.43%	78.57%	88.89%	

ตารางที่ 4.5 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM)

Accuracy: 75.00% (Support Vector Machine : SVM)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	23	4	3	3	0	69.70%
pred. Transportation	0	10	0	2	0	83.33%
pred. Attraction	0	0	0	0	0	0.00%
pred. Location	0	0	4	9	3	56.25%
pred. Weather	0	0	0	0	15	100.00%
Recall	100.00%	71.43%	0.00%	64.29%	83.33%	

ข้อมูลชุดที่ 2: ข้อมูลของเว็บเซอร์วิสที่ผ่านการหาความสัมพันธ์ของคำในออนโทโลยี

ตารางที่ 4.6 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคมีนส์ (K-means)

Accuracy: 64.47% (K-means)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	18	1	5	0	0	75.00%
pred. Transportation	4	6	1	4	0	40.00%
pred. Attraction	1	1	1	4	0	14.29%
pred. Location	0	6	0	6	0	50.00%
pred. Weather	0	0	0	0	18	100.00%
Recall	78.26%	42.86%	14.29%	42.86%	100.00%	

ตารางที่ 4.7 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB)

Accuracy: 75.18% (Naive Bayes : NB)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	22	3	3	5	1	64.71%
pred. Transportation	1	11	1	2	0	73.33%
pred. Attraction	0	0	1	0	0	100.00%
pred. Location	0	0	2	6	0	75.00%
pred. Weather	0	0	0	1	17	94.44%
Recall	95.65%	78.57%	14.29%	42.86%	94.44%	

ตารางที่ 4.8 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN)

Accuracy: 84.29% (k-Nearest Neighbor : k-NN)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	23	2	3	2	0	76.67%
pred. Transportation	0	12	0	2	0	85.71%
pred. Attraction	0	0	3	1	0	75.00%
pred. Location	0	0	1	8	0	88.89%
pred. Weather	0	0	0	1	18	94.74%
Recall	100.00%	85.71%	42.86%	57.14%	100.00%	

ตารางที่ 4.9 ผลค่าความแม่นยำและค่าความระลึกโดยเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT)

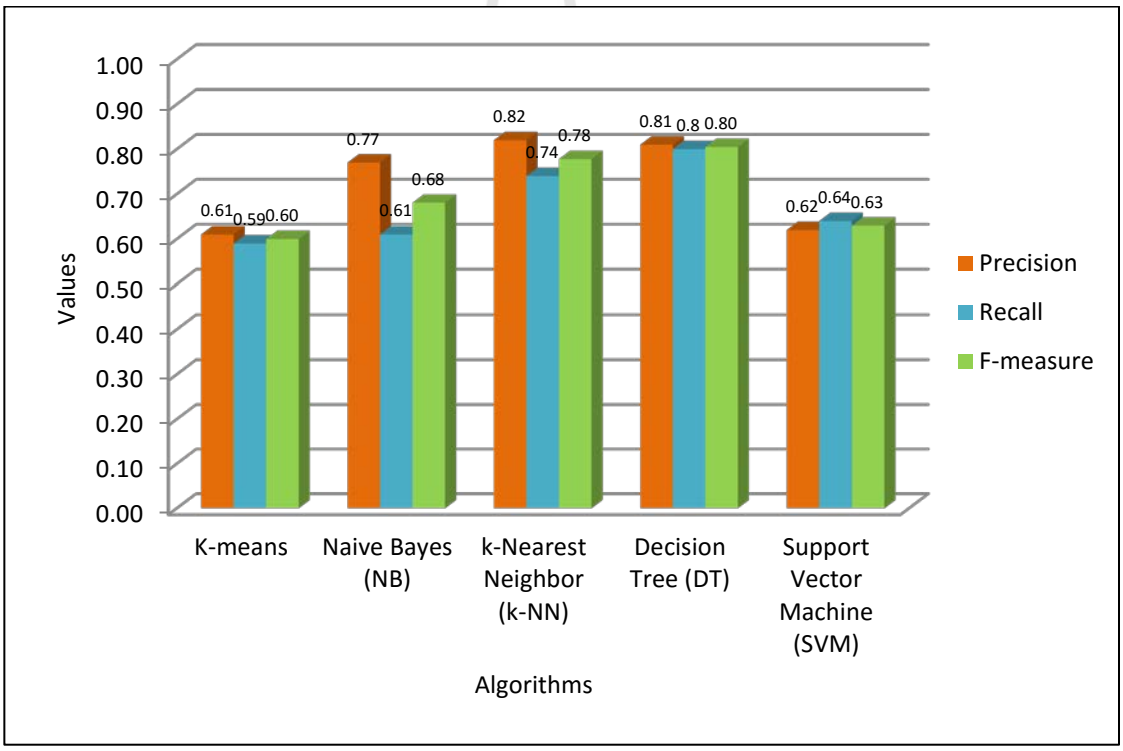
Accuracy: 96.07% (Decision Tree : DT)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	23	1	0	0	0	95.83%
pred. Transportation	0	13	0	0	0	100.00%
pred. Attraction	0	0	7	1	0	87.50%
pred. Location	0	0	0	13	1	92.86%
pred. Weather	0	0	0	0	17	100.00%
Recall	100.00%	92.86%	100.00%	92.86%	94.44%	

ตารางที่ 4.10 เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM)

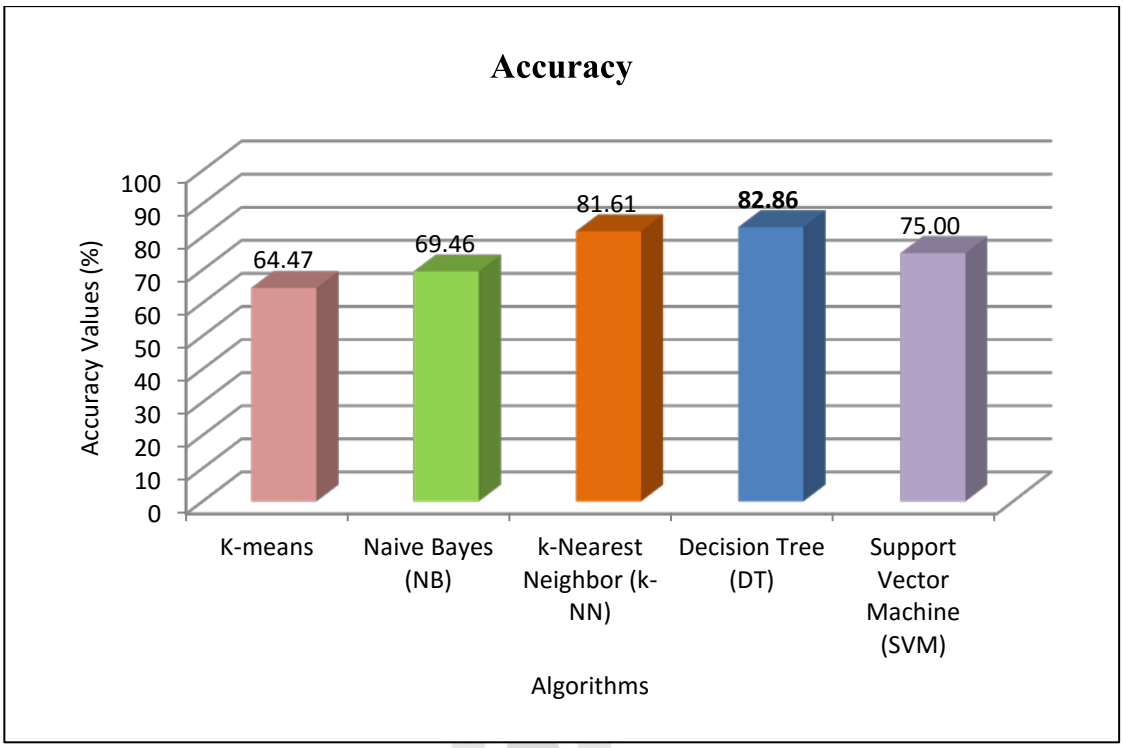
Accuracy: 84.46% (Support Vector Machine : SVM)						
	true Accommodation	true Transportation	true Attraction	true Location	true Weather	Precision
pred. Accommodation	23	1	3	1	0	82.14%
pred. Transportation	0	12	1	2	0	80.00%
pred. Attraction	0	0	0	0	0	0.00%
pred. Location	0	1	3	11	0	73.33%
pred. Weather	0	0	0	0	18	100.00%
Recall	100.00%	85.71%	0.00%	78.57%	100.00%	

ตารางที่ 4.11 สรุปผลการวัดประสิทธิภาพการจัดกลุ่มเว็บเซอร์วิส แสดงค่าความแม่นยำ ค่าความระลึก ค่าเอฟเมเซอร์ และค่าความถูกต้อง ของชุดข้อมูลในแบบที่ 1 (ข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำ)

เทคนิควิธี	Precision	Recall	F-measure	Accuracy (%)
K-means	0.61	0.59	0.60	64.47
Naive Bayes (NB)	0.77	0.61	0.68	69.46
k-Nearest Neighbor (k-NN)	0.82	0.74	0.78	81.61
Decision Tree (DT)	**0.81**	**0.80**	**0.80**	**82.86**
Support Vector Machine (SVM)	0.62	0.64	0.63	75.00



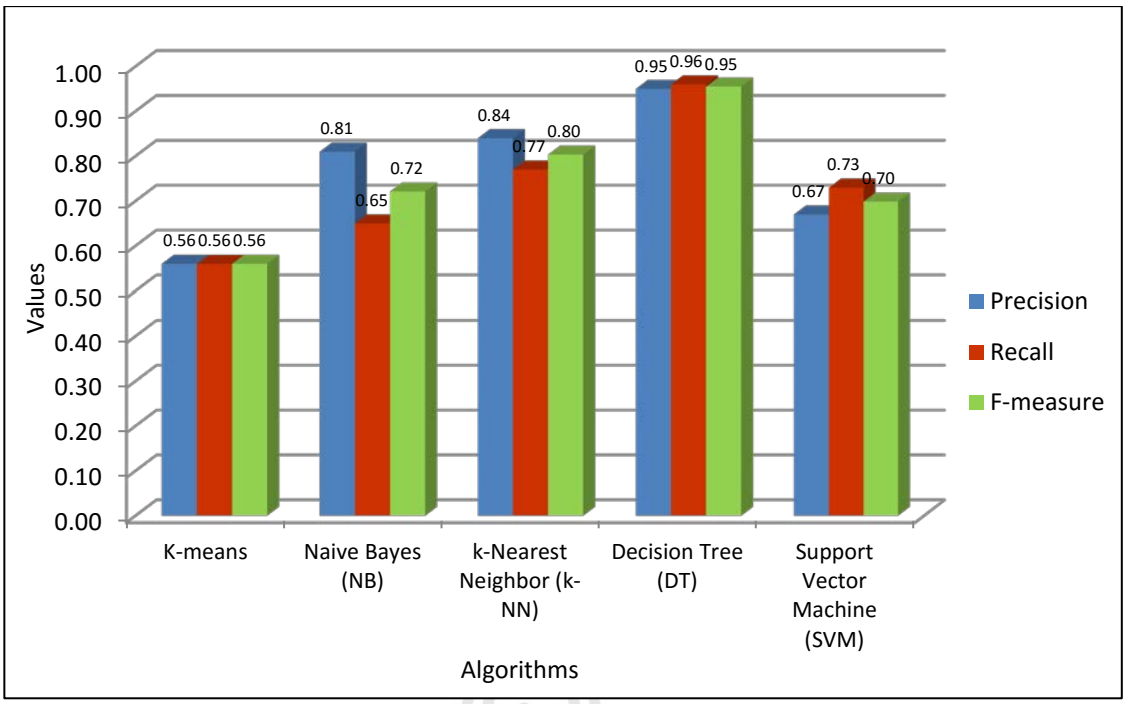
รูปที่ 4.1 แผนภูมิแสดงค่าความแม่นยำ ค่าความระลึก และค่าเอฟเมเซอร์ของชุดข้อมูลในแบบที่ 1



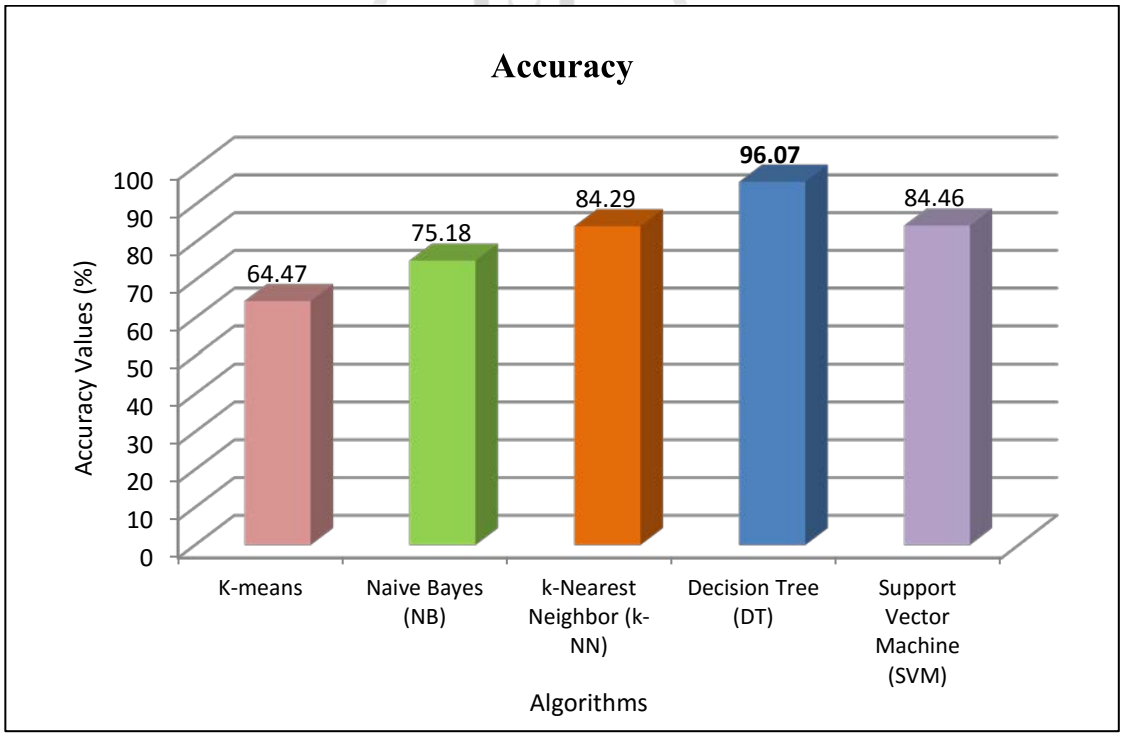
รูปที่ 4.2 แผนภูมิแสดงค่าความถูกต้องของชุดข้อมูลในแบบที่ 1

ตารางที่ 4.12 สรุปผลการวัดประสิทธิภาพการจัดกลุ่มเว็บเซอร์วิส แสดงค่าความแม่นยำ ค่าความระลึก ค่าเอฟเมเซอร์ และค่าความถูกต้อง ของชุดข้อมูลในแบบที่ 2 (ข้อมูลของเว็บเซอร์วิสที่ผ่านการหาความสัมพันธ์ของคำเชิงความหมายในออนโทโลยี)

เทคนิควิธี	Precision	Recall	F-measure	Accuracy (%)
K-means	0.56	0.56	0.56	64.47
Naive Bayes (NB)	0.81	0.65	0.72	75.18
k-Nearest Neighbor (k-NN)	0.84	0.77	0.80	84.29
Decision Tree (DT)	**0.95**	**0.96**	**0.95**	**96.07**
Support Vector Machine (SVM)	0.67	0.73	0.70	84.46



รูปที่ 4.3 แผนภูมิแสดงค่าความแม่นยำ ค่าความระลึก และค่าเอฟเมเซอร์ของชุดข้อมูลในแบบที่ 2

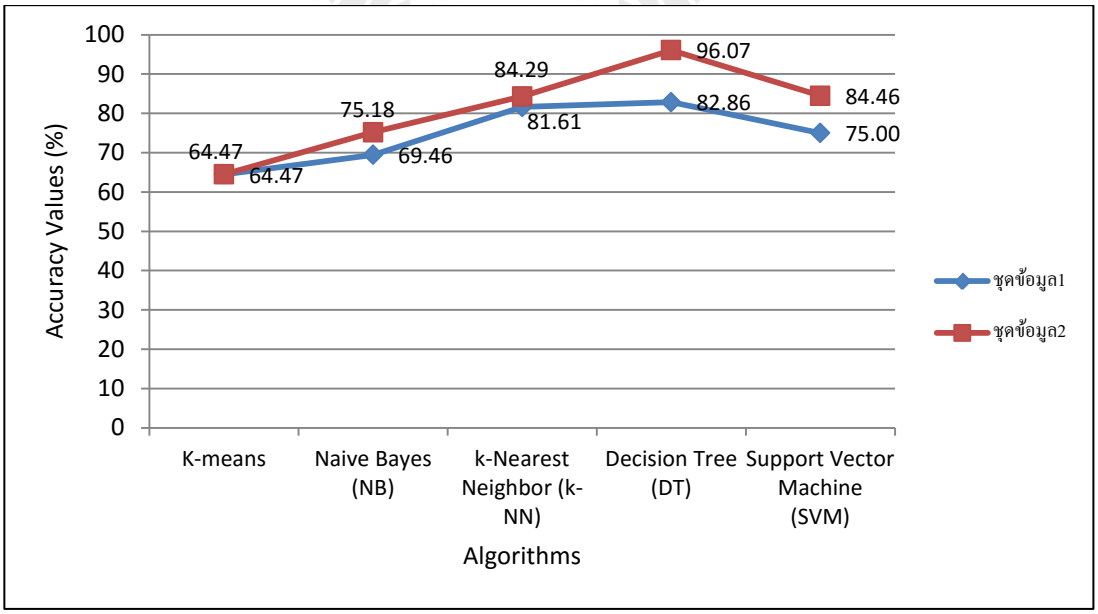


รูปที่ 4.4 แผนภูมิแสดงค่าความถูกต้องของชุดข้อมูลในแบบที่ 2

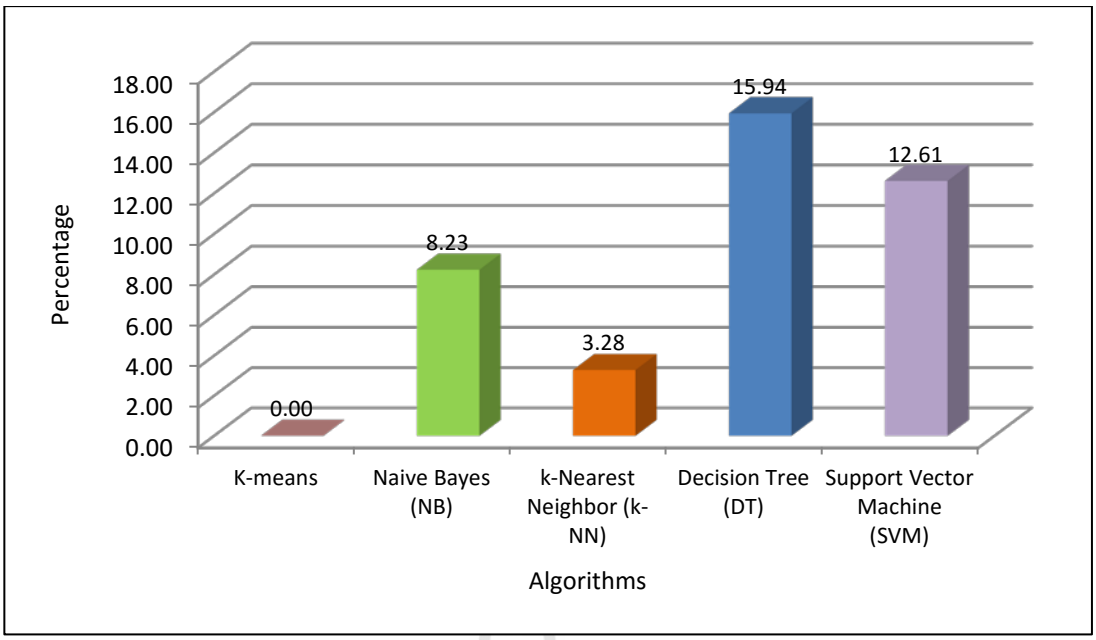
จากผลการทดลองการจัดกลุ่มเว็บเซอร์วิสพบว่า ในชุดข้อมูลแบบที่ 1 เทคนิควิธีที่ให้ค่าความถูกต้อง (Accuracy) สูงที่สุดคือเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) คือ 82.86% รองลงมาคือเทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) คือ 81.61% เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) คือ 75.00% เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) คือ 69.46% และเทคนิควิธีเคมีนส์ (K-means) คือ 64.47% แสดงดังรูปที่ 4.2

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% เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) คือ 75.18% และเทคนิควิธีเคมีนส์ (K-means) คือ 64.47% แสดงดังรูปที่ 4.4

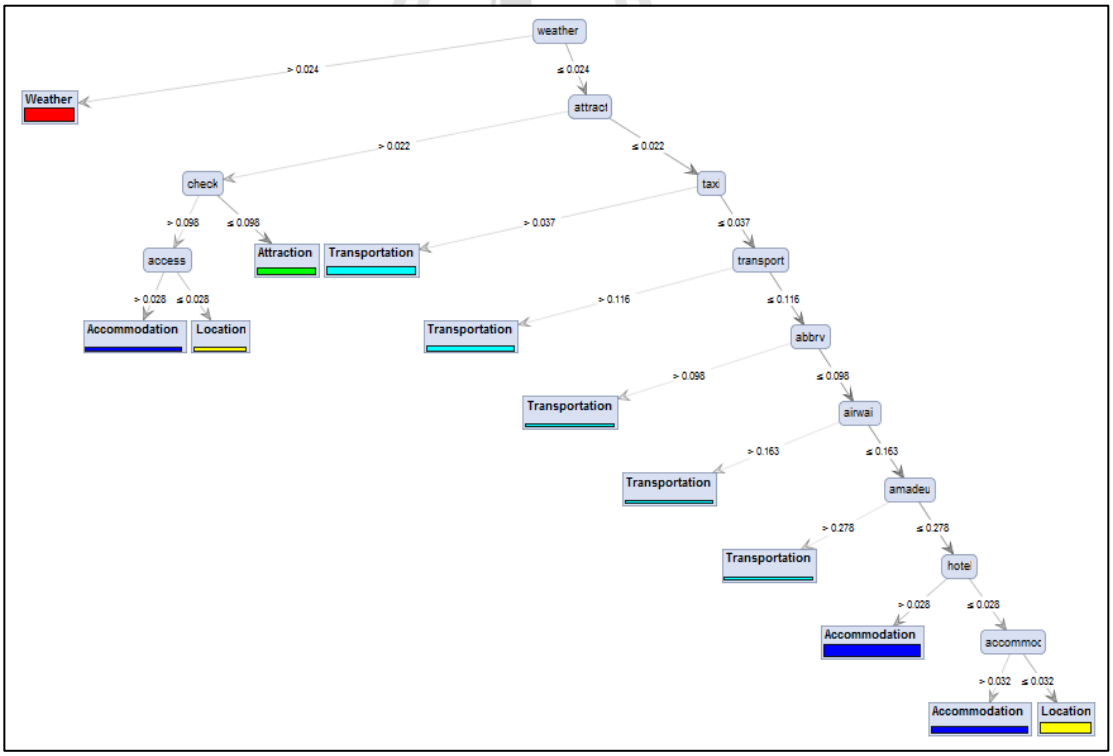
จากการเปรียบเทียบค่าความถูกต้อง (Accuracy) ของชุดข้อมูลทั้งสองชุด พบว่าชุดข้อมูลในแบบที่ 2 (ข้อมูลของเว็บเซอร์วิสที่ผ่านการหาความสัมพันธ์ของคำในออนโทโลยี) ให้ค่าความถูกต้องที่สูงกว่าชุดข้อมูลในแบบที่ 1 (ข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำ) เฉลี่ย 8.01% โดยที่เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) สูงขึ้น 15.94% เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) สูงขึ้น 12.61% เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) สูงขึ้น 8.23% เทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) สูงขึ้น 3.28% และเทคนิควิธีเคมีนส์ (K-means) มีค่าไม่เปลี่ยนแปลง แสดงดังรูปที่ 4.5 และรูปที่ 4.6



รูปที่ 4.5 แผนภูมิแสดงการเปรียบเทียบค่าความถูกต้องระหว่างข้อมูลชุดที่ 1 และข้อมูลชุดที่ 2

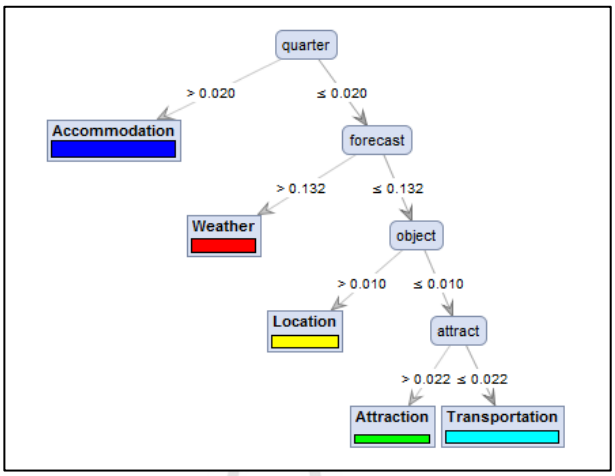


รูปที่ 4.6 แผนภูมิแสดงอัตราผลต่างของค่าความถูกต้องระหว่างข้อมูลชุดที่ 1 และข้อมูลชุดที่ 2



รูปที่ 4.7 โมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 1

รูปที่ 4.7 แสดงโมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 1 (ข้อมูลของเว็บเซอร์วิสที่ไม่ได้ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย) ซึ่งให้ค่าความถูกต้อง (Accuracy) ในการจำแนกหมวดหมู่เว็บเซอร์วิสเท่ากับ 82.86%



รูปที่ 4.8 โมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 2

รูปที่ 4.8 แสดงโมเดลการจำแนกหมวดหมู่เว็บเซอร์วิสของเทคนิควิธีต้นไม้ตัดสินใจของข้อมูลชุดที่ 2 (ข้อมูลของเว็บเซอร์วิสที่ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย) ซึ่งให้ค่าความถูกต้อง (Accuracy) ในการจำแนกหมวดหมู่เว็บเซอร์วิสเท่ากับ 96.07%

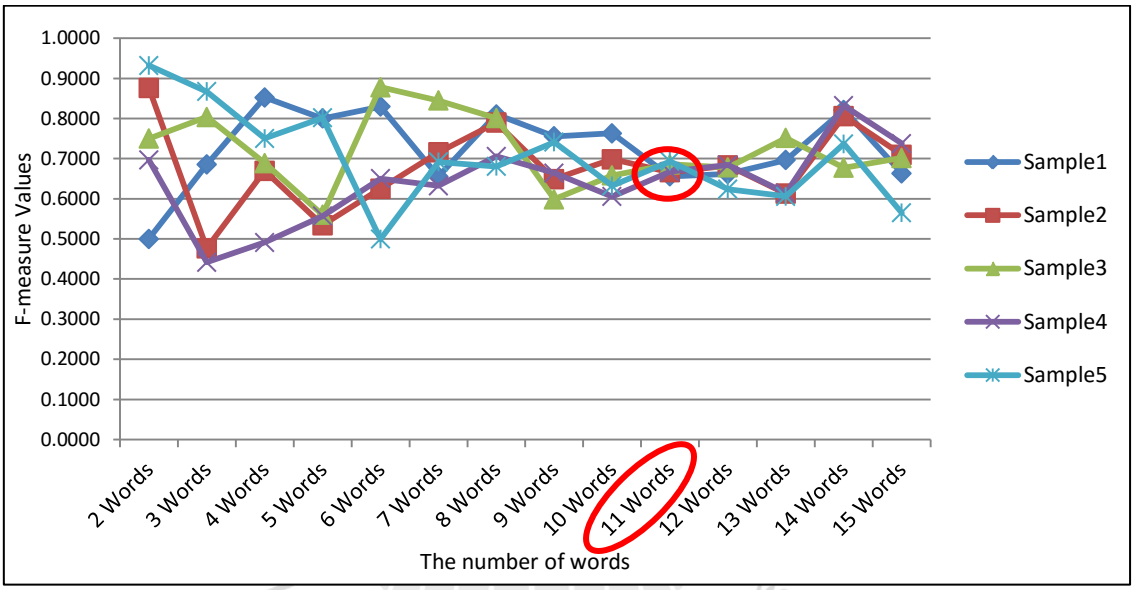
4.2 ผลการคัดเลือกเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยวและการประเมินประสิทธิภาพ

4.2.1 การหาจำนวนคำที่เหมาะสมในการประเมินประสิทธิภาพในการค้นหาเว็บเซอร์วิสเชิงความหมาย

ในการหาจำนวนคำที่เหมาะสมเพื่อใช้ในการประเมินประสิทธิภาพในการค้นหาเว็บเซอร์วิสเชิงความหมาย โดยใช้ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation : SD) เพื่อหาจำนวนคำที่เหมาะสม โดยการสุ่มคำค้นในฐานข้อมูลมาจำนวน 5 ชุดตัวอย่าง แล้วนำคำค้นที่สุ่มได้ไปค้นหาเว็บเซอร์วิส จากนั้นนำค่าเอฟเมเชอร์ (F-measure) ที่ได้จาก 5 ชุดตัวอย่างมาหาค่าเบี่ยงเบนมาตรฐาน ผลลัพธ์ค่าเอฟเมเชอร์ที่ได้และค่าเบี่ยงเบนมาตรฐานแสดงดังตารางที่ 4.13 และรูปที่ 4.9

ตารางที่ 4.13 ค่าเอฟเมเซอร์ (F-measure) และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation : SD)

	5 Words	6 Words	7 Words	8 Words	9 Words	10 Words	11 Words	12 Words	13 Words	14 Words	15 Words
Sample1	0.800000	0.829820	0.655331	0.810239	0.755570	0.762895	0.655661	0.661688	0.696532	0.821417	0.663449
Sample2	0.534091	0.625055	0.715584	0.789870	0.649006	0.698578	0.666212	0.683011	0.612882	0.806126	0.709551
Sample3	0.559885	0.877962	0.844664	0.800758	0.598789	0.658277	0.685576	0.677922	0.751638	0.677177	0.701479
Sample4	0.556723	0.650216	0.632323	0.705758	0.664105	0.605441	0.667448	0.684218	0.614388	0.832119	0.737394
Sample5	0.801732	0.499394	0.691775	0.680321	0.741069	0.632359	0.693355	0.623815	0.605842	0.736926	0.564773
SD	0.137639	0.155564	0.082914	0.059863	0.065639	0.061557	**0.015394**	0.025306	0.064992	0.065963	0.067214

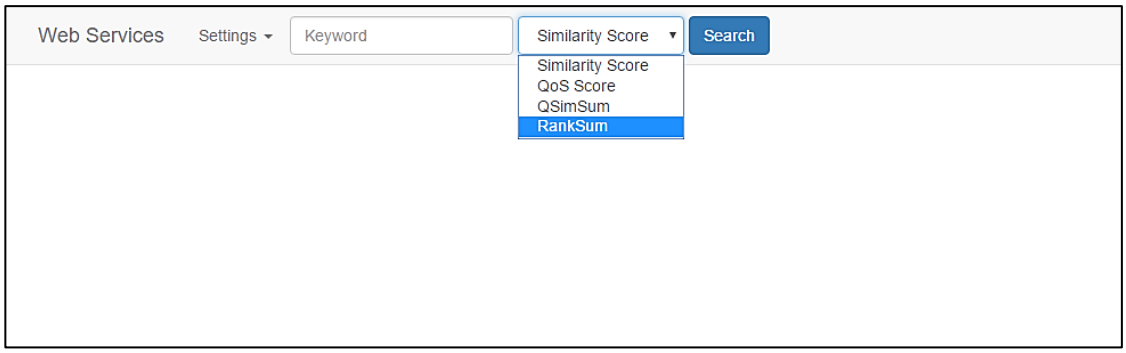


รูปที่ 4.9 แผนภูมิแสดงค่าเอฟเมเซอร์ (F-measure)

ผลการทดลองพบว่าจำนวนคำที่เหมาะสมที่จะใช้ทดสอบเพื่อประเมินประสิทธิภาพในการค้นหาเว็บเซอร์วิสเชิงความหมาย 3 อันดับแรก คือจำนวน 11 คำ ซึ่งให้ค่าเบี่ยงเบนมาตรฐาน (SD) คือ 0.015394 รองลงมาคือ 12 คำ ซึ่งให้ค่าเบี่ยงเบนมาตรฐาน (SD) คือ 0.025306 และ 8 คำ ซึ่งให้ค่าเบี่ยงเบนมาตรฐาน (SD) คือ 0.059863 ดังนั้นงานวิจัยนี้จึงเลือกคำจำนวนคำ 11 คำ เพื่อใช้ในการประเมินผลประสิทธิของระบบสืบค้นเว็บเซอร์วิสเชิงความหมาย

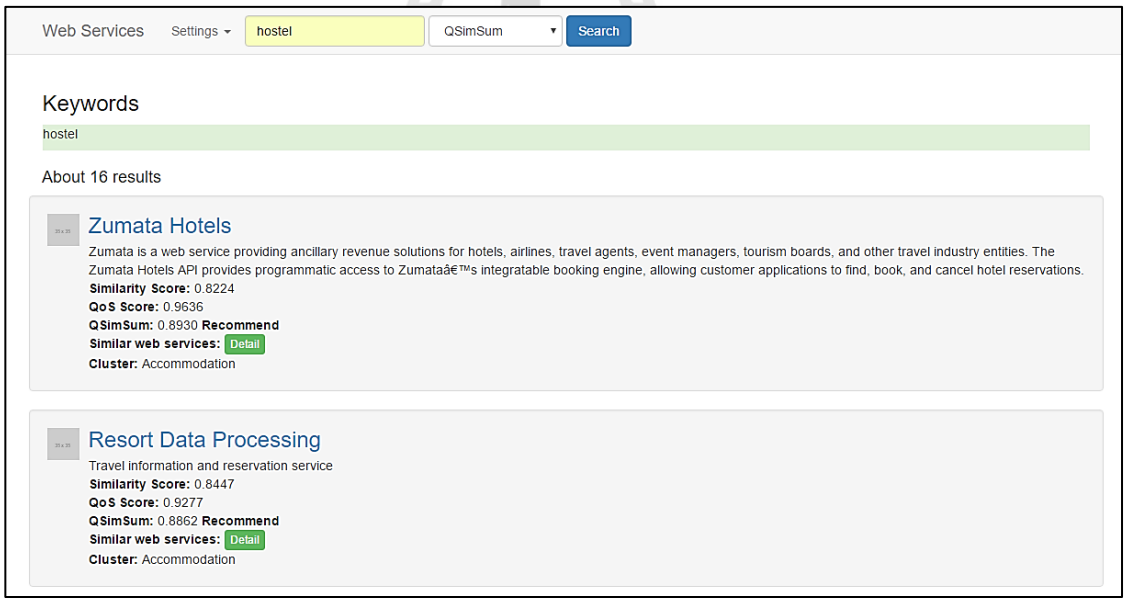
4.2.2 ผลการคัดเลือกเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยว

ตัวอย่างส่วนติดต่อกับผู้ใช้ (User Interface) แสดงดังรูปที่ 4.10 4.11 4.12 4.13 4.14 4.15 4.16 4.17 และ 4.18

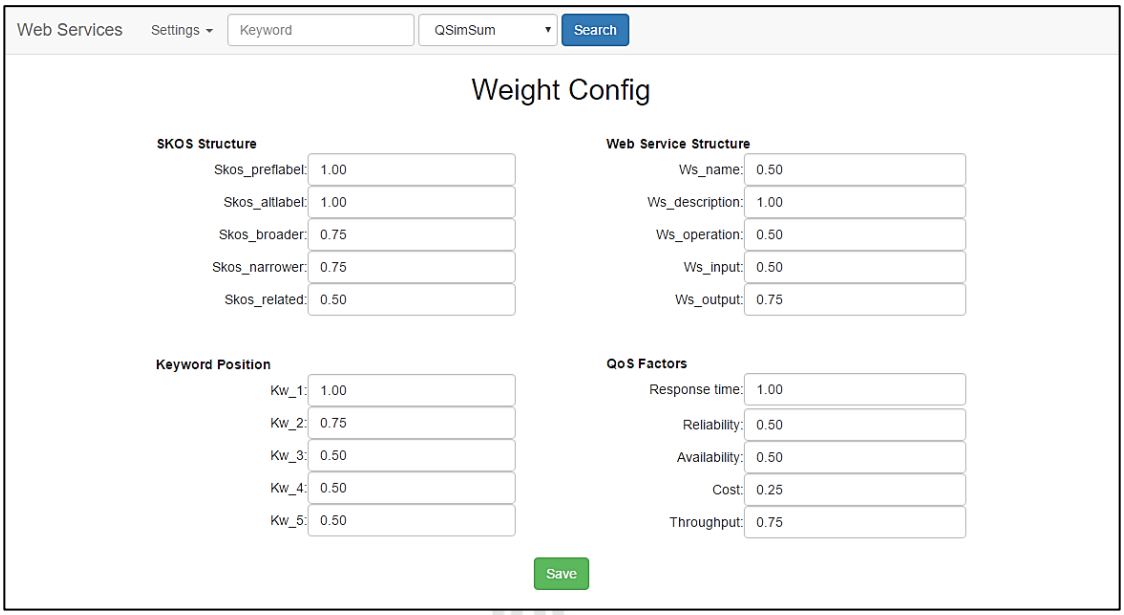


รูปที่ 4.10 หน้าแรกของระบบการค้นหาเว็บเซอร์วิสเชิงความหมาย

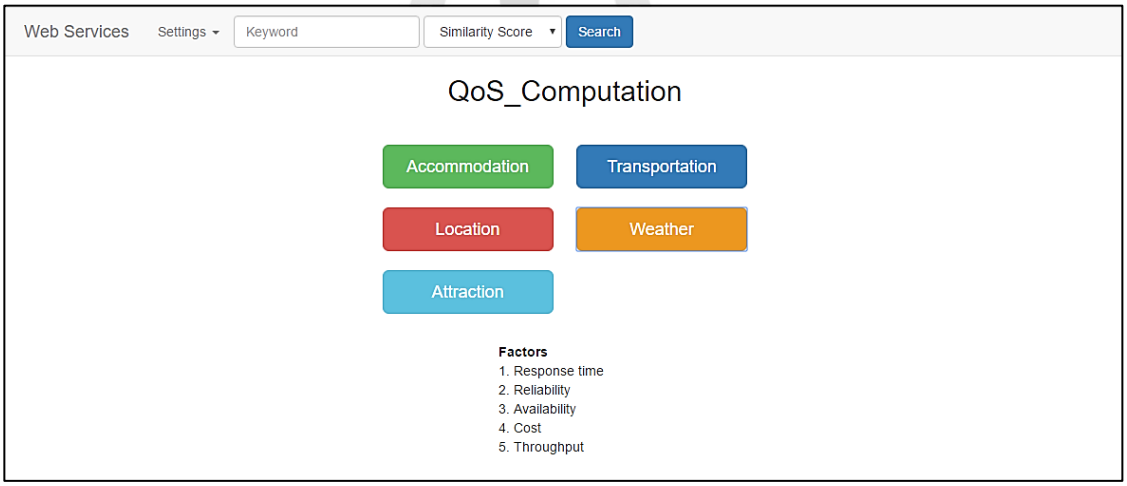
จากรูปที่ 4.10 ผู้ใช้สามารถกำหนดผลการเรียงลำดับของผลลัพธ์จากการค้นหา (User Preferences) ได้จากค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิส ค่าคะแนนความคล้ายคลึงกันเชิงความหมายระหว่างคำค้นจากผู้ใช้และข้อมูลในฐานความรู้แบบออนโทโลยี จากค่าเฉลี่ยของทั้งสองค่าข้างต้น และเรียงลำดับแบบวิธีหาผลรวม (RankSum)



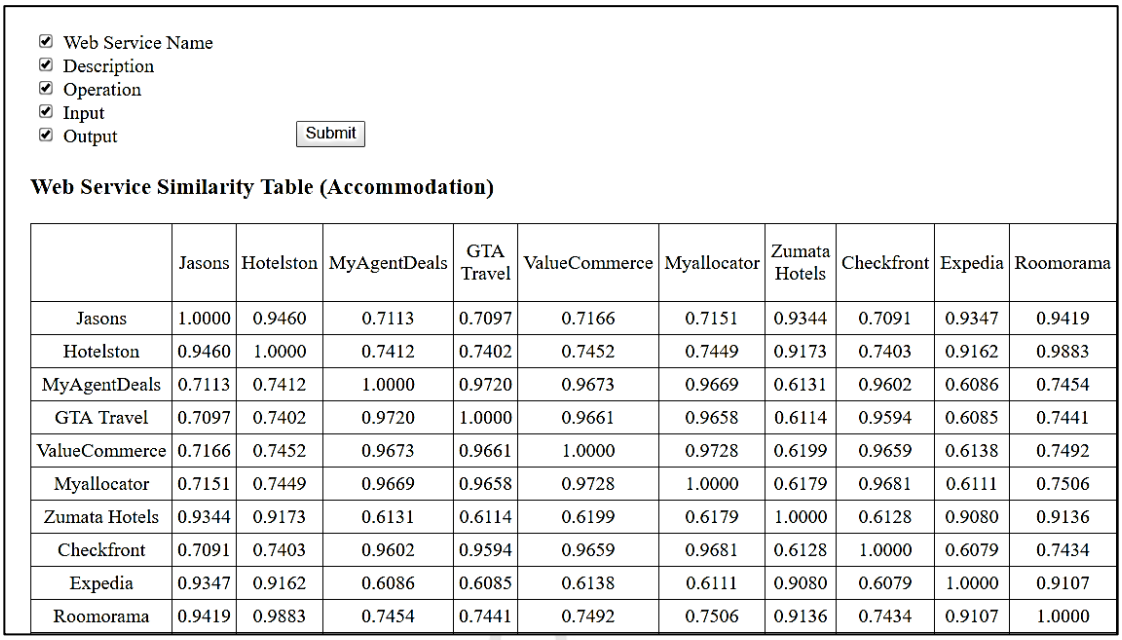
รูปที่ 4.11 ตัวอย่างผลลัพธ์เมื่อค้นหาด้วยคำค้น “hostel”



รูปที่ 4.12 หน้าการตั้งค่าน้ำหนัก (Weight Configuration)



รูปที่ 4.13 หน้าการคำนวณหาค่าคะแนนคุณภาพการบริการของแต่ละกลุ่มเว็บเซอร์วิส

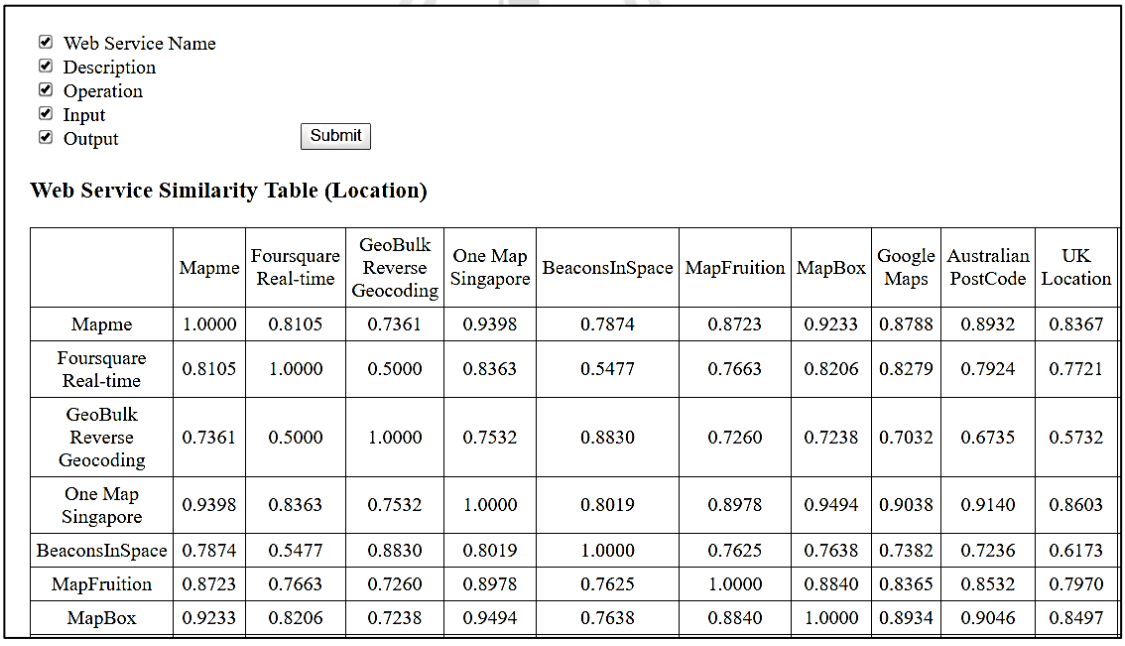


☑ Web Service Name
☑ Description
☑ Operation
☑ Input
☑ Output Submit

Web Service Similarity Table (Accommodation)

	Jasons	Hotelston	MyAgentDeals	GTA Travel	ValueCommerce	Myallocator	Zumata Hotels	Checkfront	Expedia	Roomorama
Jasons	1.0000	0.9460	0.7113	0.7097	0.7166	0.7151	0.9344	0.7091	0.9347	0.9419
Hotelston	0.9460	1.0000	0.7412	0.7402	0.7452	0.7449	0.9173	0.7403	0.9162	0.9883
MyAgentDeals	0.7113	0.7412	1.0000	0.9720	0.9673	0.9669	0.6131	0.9602	0.6086	0.7454
GTA Travel	0.7097	0.7402	0.9720	1.0000	0.9661	0.9658	0.6114	0.9594	0.6085	0.7441
ValueCommerce	0.7166	0.7452	0.9673	0.9661	1.0000	0.9728	0.6199	0.9659	0.6138	0.7492
Myallocator	0.7151	0.7449	0.9669	0.9658	0.9728	1.0000	0.6179	0.9681	0.6111	0.7506
Zumata Hotels	0.9344	0.9173	0.6131	0.6114	0.6199	0.6179	1.0000	0.6128	0.9080	0.9136
Checkfront	0.7091	0.7403	0.9602	0.9594	0.9659	0.9681	0.6128	1.0000	0.6079	0.7434
Expedia	0.9347	0.9162	0.6086	0.6085	0.6138	0.6111	0.9080	0.6079	1.0000	0.9107
Roomorama	0.9419	0.9883	0.7454	0.7441	0.7492	0.7506	0.9136	0.7434	0.9107	1.0000

รูปที่ 4.14 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Accommodation

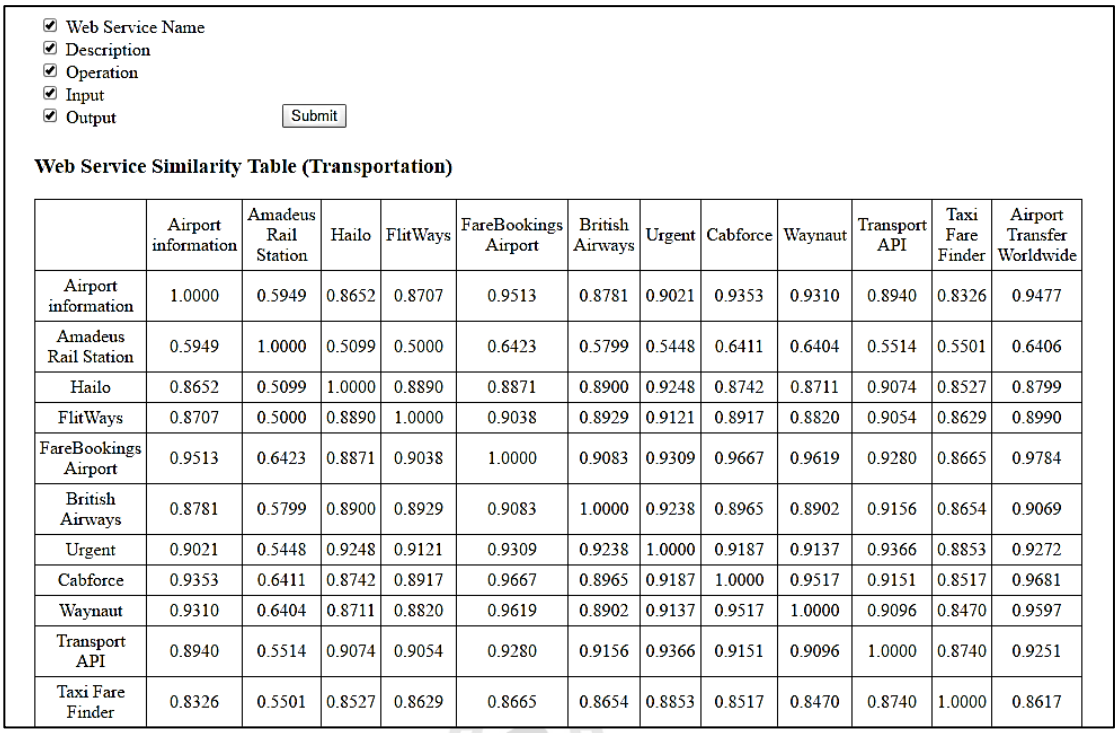


☑ Web Service Name
☑ Description
☑ Operation
☑ Input
☑ Output Submit

Web Service Similarity Table (Location)

	Mapme	Foursquare Real-time	GeoBulk Reverse Geocoding	One Map Singapore	BeaconsInSpace	MapFruition	MapBox	Google Maps	Australian PostCode	UK Location
Mapme	1.0000	0.8105	0.7361	0.9398	0.7874	0.8723	0.9233	0.8788	0.8932	0.8367
Foursquare Real-time	0.8105	1.0000	0.5000	0.8363	0.5477	0.7663	0.8206	0.8279	0.7924	0.7721
GeoBulk Reverse Geocoding	0.7361	0.5000	1.0000	0.7532	0.8830	0.7260	0.7238	0.7032	0.6735	0.5732
One Map Singapore	0.9398	0.8363	0.7532	1.0000	0.8019	0.8978	0.9494	0.9038	0.9140	0.8603
BeaconsInSpace	0.7874	0.5477	0.8830	0.8019	1.0000	0.7625	0.7638	0.7382	0.7236	0.6173
MapFruition	0.8723	0.7663	0.7260	0.8978	0.7625	1.0000	0.8840	0.8365	0.8532	0.7970
MapBox	0.9233	0.8206	0.7238	0.9494	0.7638	0.8840	1.0000	0.8934	0.9046	0.8497

รูปที่ 4.15 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Location

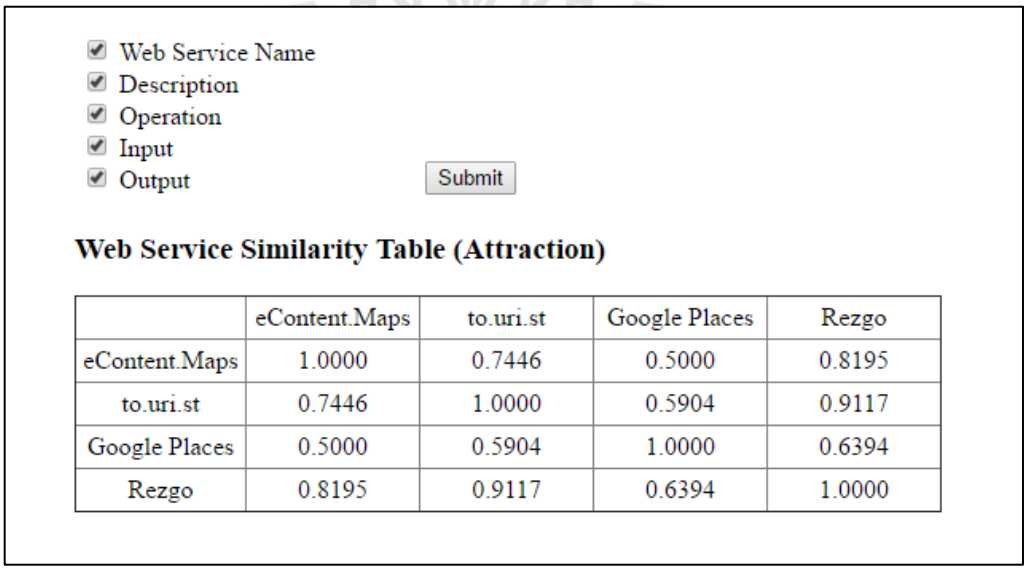


☑ Web Service Name
☑ Description
☑ Operation
☑ Input
☑ Output Submit

Web Service Similarity Table (Transportation)

	Airport information	Amadeus Rail Station	Hailo	FlitWays	FareBookings Airport	British Airways	Urgent	Cabforce	Waynaut	Transport API	Taxi Fare Finder	Airport Transfer Worldwide
Airport information	1.0000	0.5949	0.8652	0.8707	0.9513	0.8781	0.9021	0.9353	0.9310	0.8940	0.8326	0.9477
Amadeus Rail Station	0.5949	1.0000	0.5099	0.5000	0.6423	0.5799	0.5448	0.6411	0.6404	0.5514	0.5501	0.6406
Hailo	0.8652	0.5099	1.0000	0.8890	0.8871	0.8900	0.9248	0.8742	0.8711	0.9074	0.8527	0.8799
FlitWays	0.8707	0.5000	0.8890	1.0000	0.9038	0.8929	0.9121	0.8917	0.8820	0.9054	0.8629	0.8990
FareBookings Airport	0.9513	0.6423	0.8871	0.9038	1.0000	0.9083	0.9309	0.9667	0.9619	0.9280	0.8665	0.9784
British Airways	0.8781	0.5799	0.8900	0.8929	0.9083	1.0000	0.9238	0.8965	0.8902	0.9156	0.8654	0.9069
Urgent	0.9021	0.5448	0.9248	0.9121	0.9309	0.9238	1.0000	0.9187	0.9137	0.9366	0.8853	0.9272
Cabforce	0.9353	0.6411	0.8742	0.8917	0.9667	0.8965	0.9187	1.0000	0.9517	0.9151	0.8517	0.9681
Waynaut	0.9310	0.6404	0.8711	0.8820	0.9619	0.8902	0.9137	0.9517	1.0000	0.9096	0.8470	0.9597
Transport API	0.8940	0.5514	0.9074	0.9054	0.9280	0.9156	0.9366	0.9151	0.9096	1.0000	0.8740	0.9251
Taxi Fare Finder	0.8326	0.5501	0.8527	0.8629	0.8665	0.8654	0.8853	0.8517	0.8470	0.8740	1.0000	0.8617

รูปที่ 4.16 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Transportation

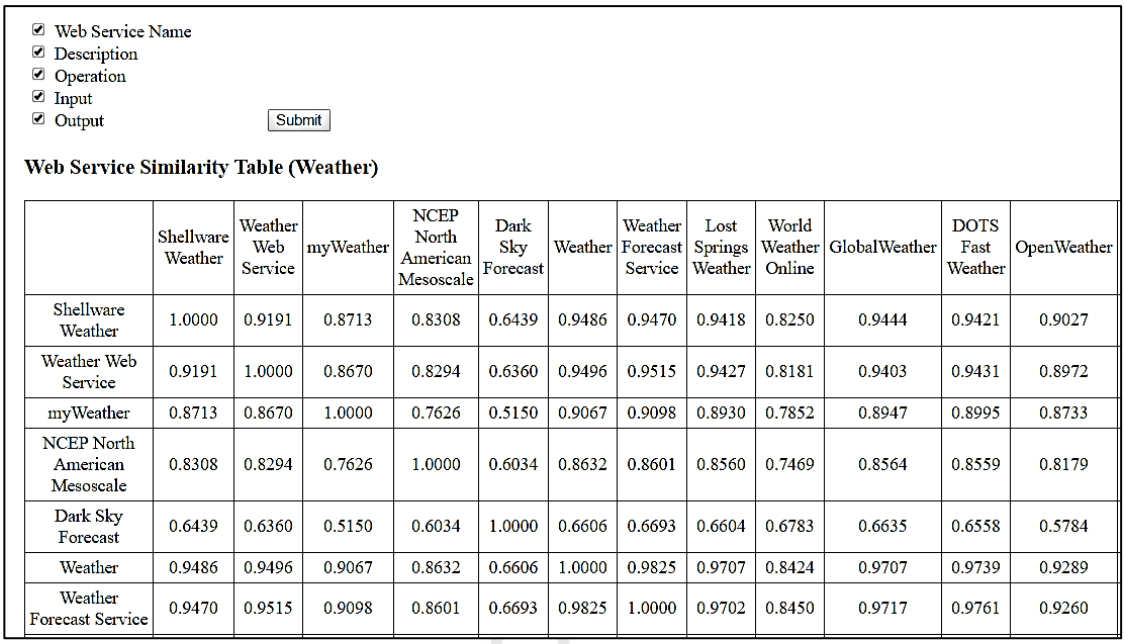


☑ Web Service Name
☑ Description
☑ Operation
☑ Input
☑ Output Submit

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

รูปที่ 4.17 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Attraction



☑ Web Service Name
☑ Description
☑ Operation
☑ Input
☑ Output Submit

Web Service Similarity Table (Weather)

	Shellware Weather	Weather Web Service	myWeather	NCEP North American Mesoscale	Dark Sky Forecast	Weather	Weather Forecast Service	Lost Springs Weather	World Weather Online	GlobalWeather	DOTS Fast Weather	OpenWeather
Shellware Weather	1.0000	0.9191	0.8713	0.8308	0.6439	0.9486	0.9470	0.9418	0.8250	0.9444	0.9421	0.9027
Weather Web Service	0.9191	1.0000	0.8670	0.8294	0.6360	0.9496	0.9515	0.9427	0.8181	0.9403	0.9431	0.8972
myWeather	0.8713	0.8670	1.0000	0.7626	0.5150	0.9067	0.9098	0.8930	0.7852	0.8947	0.8995	0.8733
NCEP North American Mesoscale	0.8308	0.8294	0.7626	1.0000	0.6034	0.8632	0.8601	0.8560	0.7469	0.8564	0.8559	0.8179
Dark Sky Forecast	0.6439	0.6360	0.5150	0.6034	1.0000	0.6606	0.6693	0.6604	0.6783	0.6635	0.6558	0.5784
Weather	0.9486	0.9496	0.9067	0.8632	0.6606	1.0000	0.9825	0.9707	0.8424	0.9707	0.9739	0.9289
Weather Forecast Service	0.9470	0.9515	0.9098	0.8601	0.6693	0.9825	1.0000	0.9702	0.8450	0.9717	0.9761	0.9260

รูปที่ 4.18 หน้าแสดงตารางค่าความคล้ายคลึงกันของเว็บเซอร์วิสในกลุ่ม Weather

4.2.3 ผลการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการ

1) ผลการประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการโดยการเปรียบเทียบผลลัพธ์ที่ได้ระหว่างผู้เชี่ยวชาญและระบบสืบค้น

จากหัวข้อที่ 4.2.1 และตารางที่ 4.13 จะพบว่าจำนวนคำที่เหมาะสมที่จะใช้ในการทดสอบเพื่อประเมินผลการสืบค้นเว็บเซอร์วิสอยู่ที่จำนวน 11 คำ ซึ่งผู้วิจัยได้สุ่มเลือกคำสำคัญจากคลังคำในฐานความรู้แบบออนโทโลยี แสดงดังตารางที่ 4.14

ตารางที่ 4.14 ตัวอย่างคำสำคัญที่ใช้ในการประเมินผลการสืบค้นเว็บเซอร์วิส

หมวดหมู่เว็บเซอร์วิส	คำสำคัญ
Accommodation	hostel, booking, resort
Transportation	flight, airway
Attraction	tour, tourism
Location	latitude, longitude
Weather	temperature, celsius

ในการประเมินประสิทธิภาพของการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น โดยใช้ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ตามสมการที่ 3.21 สมการที่ 3.22 สมการที่ 3.23 และตามสมการที่ 3.24 (ในบทที่ 3)

ตารางที่ 4.15 จำนวนเว็บเซอร์วิสที่ถูกค้นพบเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้นแยกตามคำค้น/คำสำคัญ และแสดงค่า Truth Table

คำค้น/คำสำคัญ	จำนวนเว็บเซอร์วิสที่ค้นพบ		Truth Table			
	ผู้เชี่ยวชาญ	ระบบสืบค้น	TP	TN	FP	FN
hostel	22	18	18	54	0	4
booking	26	23	22	49	1	4
resort	23	19	17	51	2	6
flight	12	11	4	57	7	8
airway	8	9	6	65	3	2
tour	12	13	10	61	3	2
tourism	12	13	11	62	2	1
latitude	24	24	20	48	4	4
longitude	24	20	18	50	2	6
temperature	18	18	18	58	0	0
celsius	18	18	18	58	0	0

ตารางที่ 4.16 ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น

คำค้น/คำสำคัญ	ค่าความแม่นยำ (Precision) TP / (TP+FP)	ค่าความระลึก (Recall) TP / (TP+FN)
hostel	18 / (18+0) = **1.0000**	18 / (18+4) = **0.8182**
booking	22 / (22+1) = **0.9565**	22 / (22+4) = **0.8462**
resort	17 / (17+2) = **0.8947**	17 / (17+6) = **0.7391**
flight	4 / (4+7) = **0.3636**	4 / (4+8) = **0.3333**

ตารางที่ 4.16 ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น (ต่อ)

คำค้น/คำสำคัญ	ค่าความแม่นยำ (Precision) TP / (TP+FP)	ค่าความระลึก (Recall) TP / (TP+FN)
airway	6 / (6+3) = **0.6667**	6 / (6+2) = **0.7500**
tour	10 / (10+3) = **0.7692**	10 / (10+2) = **0.8333**
tourism	11 / (11+2) = **0.8462**	11 / (11+1) = **0.9167**
latitude	20 / (20+4) = **0.8333**	20 / (20+4) = **0.8333**
longitude	18 / (18+2) = **0.9000**	18 / (18+6) = **0.7500**
temperature	18 / (18+0) = **1.0000**	18 / (18+0) = **1.0000**
celsius	18 / (18+0) = **1.0000**	18 / (18+0) = **1.0000**
ค่าเฉลี่ย	**0.8391**	**0.8018**

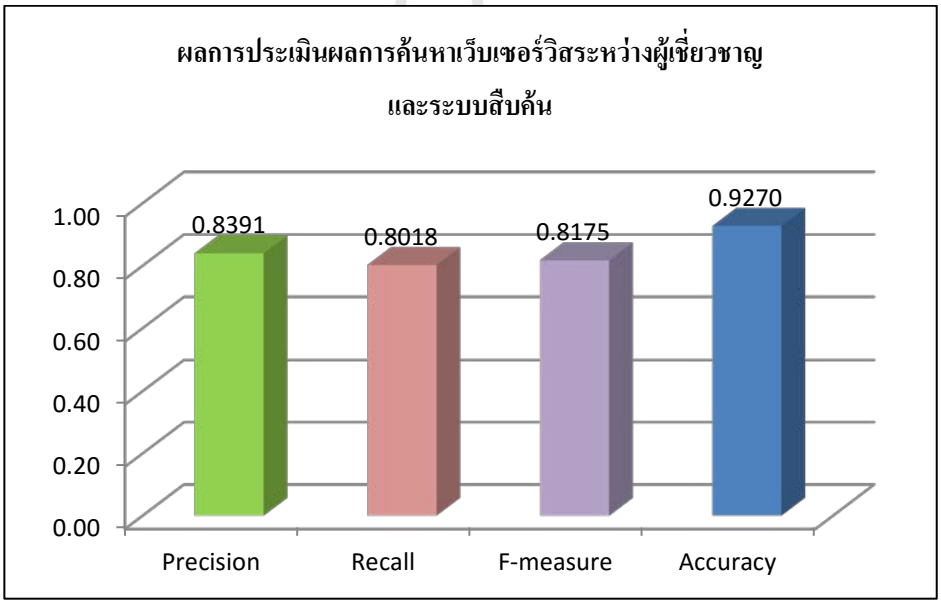
ตารางที่ 4.17 ค่าเอฟเมเซอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น

คำค้น/คำสำคัญ	ค่าเอฟเมเซอร์ (F-measure) 2 * (P * R) / (P+R)	ค่าความถูกต้อง (Accuracy) (TP+TN) / (TP+TN+FP+FN)
hostel	2 * (1.00*0.82) / (1.00+0.82) = **0.90**	(18+54) / (18+54+0+4) = **0.95**
booking	2 * (0.96*0.85) / (0.96+0.85) = **0.90**	(22+49) / (22+49+1+4) = **0.93**
resort	2 * (0.89*0.74) / (0.89+0.74) = **0.81**	(17+51) / (17+51+2+6) = **0.89**
flight	2 * (0.36*0.33) / (0.36+0.33) = **0.35**	(4+57) / (4+57+7+8) = **0.80**
airway	2 * (0.67*0.75) / (0.67+0.75) = **0.71**	(6+65) / (6+65+3+2) = **0.93**
tour	2 * (0.77*0.83) / (0.77+0.83) = **0.80**	(10+61) / (10+61+3+2) = **0.93**
tourism	2 * (0.85*0.92) / (0.85+0.92) = **0.88**	(11+62) / (11+62+2+1) = **0.96**
latitude	2 * (0.83*0.83) / (0.83+0.83) = **0.83**	(20+48) / (20+48+4+4) = **0.89**
longitude	2 * (0.90*0.75) / (0.90+0.75) = **0.82**	(18+50) / (18+50+2+6) = **0.89**
temperature	2 * (1.00*1.00) / (1.00+1.00) = **1.00**	(18+58) / (18+58+0+0) = **1.00**
celsius	2 * (1.00*1.00) / (1.00+1.00) = **1.00**	(18+58) / (18+58+0+0) = **1.00**
ค่าเฉลี่ย	**0.8175**	**0.9270**

ค่าผลลัพธ์จากการทดสอบประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการโดยเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้น สามารถสรุปแสดงได้ดังตารางที่ 4.18 และรูปที่ 4.19

ตารางที่ 4.18 ผลการประเมินผลการค้นหาเว็บเซอร์วิสระหว่างผู้เชี่ยวชาญและระบบสืบค้น

วิธีการ	ค่าผลลัพธ์
Precision	0.8391
Recall	0.8018
F-measure	0.8175
Accuracy	0.9270



รูปที่ 4.19 ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายระหว่างผู้เชี่ยวชาญและระบบสืบค้น

ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายโดยเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้น ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.8391 ค่าความระลึก (Recall) เท่ากับ 0.8018 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.8175 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.9270

2) ผลการประเมินผลการค้นหาเว็บเซอร์วิสเชิงความหมายตามคุณภาพการบริการโดยใช้ไดนามิกเกณฑ์

การประเมินประสิทธิภาพของการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์ประเมินโดยใช้ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) แสดงผลดังตารางที่ 4.19 4.20 4.21 และ 4.22

ตารางที่ 4.19 จำนวนเว็บเซอร์วิสที่ถูกค้นพบโดยใช้ไดนามิกเกณฑ์ และแสดงค่า Truth Table

คำค้น/คำสำคัญ	จำนวนเว็บเซอร์วิสที่ค้นพบ		Truth Table			
	Matching	Semantic	TP	TN	FP	FN
hostel	2	15	2	61	13	0
booking	20	23	20	53	3	0
resort	2	14	2	62	12	0
flight	6	8	6	68	2	0
airway	1	3	1	73	2	0
tour	13	13	13	63	0	0
tourism	4	5	4	71	1	0
latitude	16	24	16	52	8	0
longitude	16	20	16	56	4	0
temperature	3	13	3	63	10	0
celsius	3	6	3	70	3	0

ตารางที่ 4.20 ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์

คำค้น/คำสำคัญ	ค่าความแม่นยำ (Precision) TP / (TP+FP)	ค่าความระลึก (Recall) TP / (TP+FN)
hostel	2 / (2+13) = **0.1333**	2 / (2+0) = **1.0000**
booking	20 / (20+3) = **0.8696**	20 / (20+0) = **1.0000**
resort	2 / (2+12) = **0.1429**	2 / (2+0) = **1.0000**
flight	6 / (6+2) = **0.7500**	6 / (6+0) = **1.0000**

ตารางที่ 4.20 ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์ (ต่อ)

คำค้น/คำสำคัญ	ค่าความแม่นยำ (Precision) TP / (TP+FP)	ค่าความระลึก (Recall) TP / (TP+FN)
airway	1 / (1+2) = **0.3333**	1 / (1+0) = **1.0000**
tour	13 / (13+0) = **1.0000**	13 / (13+0) = **1.0000**
tourism	4 / (4+1) = **0.8000**	4 / (4+0) = **1.0000**
latitude	16 / (16+8) = **0.6667**	16 / (16+0) = **1.0000**
longitude	16 / (16+4) = **0.8000**	16 / (16+0) = **1.0000**
temperature	3 / (3+10) = **0.2308**	3 / (3+0) = **1.0000**
celsius	3 / (3+3) = **0.5000**	3 / (3+0) = **1.0000**
ค่าเฉลี่ย	**0.5660**	**1.0000**

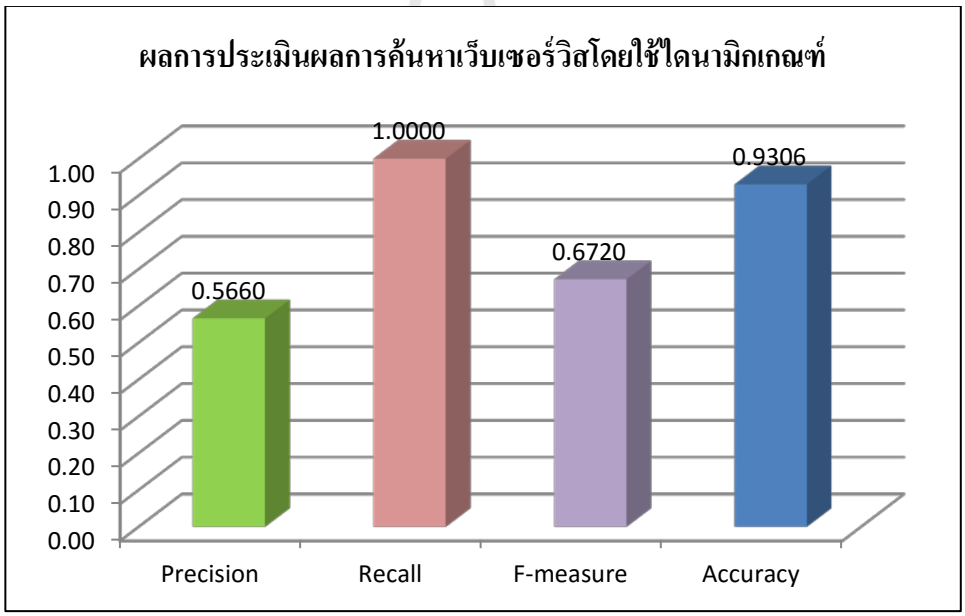
ตารางที่ 4.21 ค่าเอฟเมเชอร์ (F-measure) และค่าความถูกต้อง (Accuracy) ของการประเมินประสิทธิภาพการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์

คำค้น/คำสำคัญ	ค่าเอฟเมเชอร์ (F-measure) 2 * (P * R) / (P+R)	ค่าความถูกต้อง (Accuracy) (TP+TN) / (TP+TN+FP+FN)
hostel	2 * (0.13*1.00) / (0.13+1.00) = **0.24**	(2+61) / (2+61+13+0) = **0.83**
booking	2 * (0.87*1.00) / (0.87+1.00) = **0.93**	(20+53) / (20+53+3+0) = **0.96**
resort	2 * (0.14*1.00) / (0.14+1.00) = **0.25**	(2+62) / (2+62+12+0) = **0.84**
flight	2 * (0.75*1.00) / (0.75+1.00) = **0.86**	(6+68) / (6+68+2+0) = **0.97**
airway	2 * (0.33*1.00) / (0.33+1.00) = **0.50**	(1+73) / (1+73+2+0) = **0.97**
tour	2 * (1.00*1.00) / (1.00+1.00) = **1.00**	(13+63) / (13+63+0+0) = **1.00**
tourism	2 * (0.80*1.00) / (0.80+1.00) = **0.89**	(4+71) / (4+71+1+0) = **0.99**
latitude	2 * (0.67*1.00) / (0.67+1.00) = **0.80**	(16+52) / (16+52+8+0) = **0.89**
longitude	2 * (0.80*1.00) / (0.80+1.00) = **0.89**	(16+56) / (16+56+4+0) = **0.95**
temperature	2 * (0.23*1.00) / (0.23+1.00) = **0.38**	(3+63) / (3+63+10+0) = **0.87**
celsius	2 * (0.50*1.00) / (0.50+1.00) = **0.67**	(3+70) / (3+70+3+0) = **0.96**
ค่าเฉลี่ย	**0.6720**	**0.9306**

ค่าผลลัพธ์จากการทดสอบประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการโดยใช้ไดนามิกเกณฑ์ สามารถสรุปแสดงได้ดังตารางที่ 4.22 และรูปที่ 4.20

ตารางที่ 4.22 ผลการประเมินผลการค้นหาเว็บเซอร์วิสโดยใช้ไดนามิกเกณฑ์

วิธีการ	ค่าผลลัพธ์
Precision	0.5660
Recall	1.0000
F-measure	0.6720
Accuracy	0.9306



รูปที่ 4.20 ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายโดยใช้ไดนามิกเกณฑ์

ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายโดยใช้ไดนามิกเกณฑ์ ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.5660 ค่าความระลึก (Recall) เท่ากับ 1.0000 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.6720 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.9306

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

ในบทนี้นำเสนอการสรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะเพิ่มเติมสำหรับการวิจัยในครั้งต่อไป โดยมีรายละเอียดดังนี้

5.1 สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) เพื่อออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว และ 2) เพื่อวิเคราะห์และเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่เกี่ยวข้องกับการบริการด้านการท่องเที่ยว โดยอาศัยหลักการ แนวคิด ทฤษฎี และเทคนิควิธีการต่าง ๆ ซึ่งผลการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว สามารถสรุปได้ดังรายละเอียดต่อไปนี้

5.1.1 สรุปผลการจำแนกหมวดหมู่ของเว็บเซอร์วิส

การจำแนกหมวดหมู่ของเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวโดยใช้วิธีการทำเหมืองข้อมูล (Data Mining) ได้มีการทดลองในหลายเทคนิควิธีเพื่อเปรียบเทียบว่าเทคนิคใดจะมีประสิทธิภาพและเหมาะสมกับข้อมูลเว็บเซอร์วิสที่ให้บริการที่เกี่ยวข้องด้านการท่องเที่ยวมากที่สุด ผลการประเมินผลและวัดประสิทธิภาพพบว่า เทคนิควิธีที่ให้ค่าความถูกต้อง (Accuracy) สูงที่สุดในการจำแนกหมวดหมู่ของเว็บเซอร์วิสคือเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) มีค่าเท่ากับ 96.07% รองลงมาคือเทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) มีค่าเท่ากับ 84.46% เทคนิควิธีเคเนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) มีค่าเท่ากับ 84.29% เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) มีค่าเท่ากับ 75.18% และเทคนิควิธีเคมีนส์ (K-means) มีค่าเท่ากับ 64.47% ตามลำดับ

5.1.2 สรุปผลการออกแบบและพัฒนาระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยว

1) ผลการพัฒนาฐานความรู้ออนโทโลยีเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว

การออกแบบและพัฒนาฐานความรู้ออนโทโลยีเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวได้ประยุกต์จากออนโทโลยีของเว็บเซอร์วิสที่เป็นมาตรฐานของดับเบิ้ลยูทรีซีที่มีชื่อว่า

“Web Service Modeling Ontology: WSMO” (W3C-WSMO, www, 2005) จากนั้นได้ปรับปรุงให้สอดคล้องเพื่อรองรับกับข้อมูลด้านการท่องเที่ยว มีการจัดทำดัชนีของคำศัพท์ที่เกี่ยวข้องกับการท่องเที่ยวไว้ในคลังคำโดยใช้สคอส มีการสร้างความสัมพันธ์ของคำศัพท์เป็นโครงสร้างแบบลำดับชั้นตามรูปแบบอภิธานศัพท์มาตรฐานของเวิร์ดเน็ตซึ่งเป็นฐานข้อมูลคำศัพท์ภาษาอังกฤษขนาดใหญ่ และมีการเก็บรวบรวมข้อมูลเว็บเซอร์วิสจากหลายหลายเซอร์วิสรีจิสทรีมาสร้างเป็นฐานความรู้แบบออนโทโลยีเพื่อให้สามารถอธิบายข้อมูลที่มีความเกี่ยวข้องในเชิงความหมายรองรับกับการสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายได้อย่างมีประสิทธิภาพ และเป็นต้นแบบของฐานความรู้ออนโทโลยีเว็บเซอร์วิสที่สามารถนำไปประยุกต์ใช้กับข้อมูลในโดเมนอื่น ๆ ได้

2) ผลการออกแบบและพัฒนาระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยว

งานวิจัยนี้ได้ออกแบบและพัฒนาระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยว ซึ่งมีส่วนติดต่อกับผู้ใช้ (User Interface) ผู้ใช้สามารถสืบค้นเว็บเซอร์วิสได้จริง อีกทั้งผู้ใช้ยังสามารถกำหนดคุณลักษณะต่าง ๆ (User Preferences) ในระบบสืบค้นให้เหมาะสมกับความต้องการของผู้ใช้ได้ เช่น กำหนดค่าน้ำหนักของปัจจัยคุณภาพ กำหนดค่าน้ำหนักของคำโครงสร้างสคอส หรือการแสดงผลลัพธ์ผู้ใช้สามารถเลือกได้ว่าจะเรียงลำดับผลลัพธ์จากค่าคะแนนคุณภาพ (QoS) ค่าคะแนนความคล้ายคลึง (Similarity) หรือเรียงลำดับผลลัพธ์แบบวิธีหาผลรวม (RankSum) เป็นต้น

ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยวโดยเปรียบเทียบระหว่างผู้เชี่ยวชาญและระบบสืบค้น ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.8391 ค่าความระลึก (Recall) เท่ากับ 0.8018 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.8175 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.9270

ผลการประเมินประสิทธิภาพของระบบสืบค้นเว็บเซอร์วิสเชิงความหมายตามคุณภาพการให้บริการสำหรับการท่องเที่ยวโดยใช้ไดนามิกเกณฑ์ ให้ค่าความแม่นยำ (Precision) เท่ากับ 0.5660 ค่าความระลึก (Recall) เท่ากับ 1.0000 ค่าเอฟเมเชอร์ (F-measure) เท่ากับ 0.6720 และค่าความถูกต้อง (Accuracy) เท่ากับ 0.9306

5.2 อภิปรายผลการวิจัย

งานวิจัยนี้ได้ทำการออกแบบและพัฒนาแบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว ซึ่งผลที่ได้จากการวิจัยเป็นไปตามวัตถุประสงค์ดังนี้

5.2.1 แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวหรือ SEMO ประกอบด้วย 1) การสร้างฐานความรู้โดยการออกแบบและพัฒนาออนโทโลยี ซึ่งประยุกต์จากออนโทโลยีเว็บเซอร์วิสที่ชื่อ Web Service Modeling Ontology (WSMO) (W3C-WSMO, www, 2005) ของดับเบิ้ลยูทรีซี และสร้างคลังคำโดยใช้สคอส (SKOS) 2) คุณภาพการบริการของเว็บเซอร์วิส มีการคำนวณหาค่าคะแนนคุณภาพของเว็บเซอร์วิส โดยพิจารณาจากปัจจัยคุณภาพที่สำคัญจำนวน 5 ปัจจัย ได้แก่ เวลาที่ตอบสนอง (Response Time) สภาพความพร้อมในการให้บริการ (Availability) ความเสถียรภาพหรือความน่าเชื่อถือ (Reliability) ค่าใช้จ่ายในการเรียกใช้บริการ (Cost) และความสำเร็จในการนำส่งข้อมูล (Throughput) สอดคล้องกับงานวิจัยของคาร์โดโซ และคณะ (Cardoso et al., 2004) ซูซิลาและวาดิเวล (Susila and Vadivel, 2014) มูลูกานาเดม กูพู และชานมูกา (Muruganadam, Gopu and Shanmuga, 2014) เทาและเกวเจย์ (Tao and Kwei-Jay, 2004) ที่ได้ทำการศึกษาวิจัยเกี่ยวกับคุณภาพการให้บริการของเว็บเซอร์วิส พบว่ามีการนำปัจจัยด้านคุณภาพมาใช้เป็นตัวกำหนดถึงคุณภาพการให้บริการ แต่จะแตกต่างกันไปตามบริบทและโดเมนของงานวิจัยนั้น และ 3) ระบบสืบค้นข้อมูลเว็บเซอร์วิสเชิงความหมาย ผู้ใช้สามารถสืบค้นเว็บเซอร์วิสและสามารถกำหนดคุณลักษณะต่าง ๆ ในระบบสืบค้นให้เหมาะสมกับความต้องการของผู้ใช้ได้ เช่น กำหนดค่าน้ำหนักของปัจจัยคุณภาพต่าง ๆ สอดคล้องกับงานวิจัยของเอเจาและเดริส (Ajao and Deris, 2013) แบดร์และคณะ (Badr et al, 2008) โมฮีบิ ดิง และชิ (Mohebi, Ding and Chi, 2012) ที่ทำวิจัยเกี่ยวกับการค้นหาเว็บเซอร์วิสที่ผู้ใช้สามารถกำหนดคุณลักษณะในการสืบค้นที่ผู้ใช้ต้องการได้

5.2.2 การวิเคราะห์และเปรียบเทียบเทคนิคในการจำแนกหมวดหมู่ของเว็บเซอร์วิสที่เกี่ยวข้องกับการบริการด้านการท่องเที่ยว ซึ่งข้อมูลในการเปรียบเทียบนี้เป็นข้อมูลเว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยว โดยแบ่งข้อมูลออกเป็น 2 ชุด คือชุดข้อมูลของเว็บเซอร์วิสที่ผ่านและไม่ได้ผ่านการหาความสัมพันธ์ของคำเชิงความหมาย จากนั้นทำการเปรียบเทียบการจำแนกข้อมูลเว็บเซอร์วิสระหว่างข้อมูลทั้งสองชุด โดยวิธีการทำเหมืองข้อมูล (Data Mining) โดยใช้ 5 เทคนิค ได้แก่ เทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree : DT) เทคนิควิธีเค-เนียร์เรสท์เนเบอร์ (k-Nearest Neighbor : k-NN) เทคนิควิธีเนอีฟเบย์ (Naive Bayes : NB) เทคนิควิธีซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine : SVM) และเทคนิควิธีเคมีนส์ (K-means) และ 3) การหาความคล้ายคลึงกันของเว็บเซอร์วิสที่อยู่ในกลุ่มบริการเดียวกัน ใช้วิธีการหาค่าความคล้ายคลึงตามทฤษฎีเวกเตอร์สเปซโมเดล คือการวัดความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) สอดคล้องกับงานวิจัยของแซลตัน (Salton, 1988) และไซโดรอฟและคณะ (Sidorov et al., 2014) ซึ่งในหลายงานวิจัยมักใช้กับชุดข้อมูลทดสอบสาธารณะ (Public Datasets) แต่ในงานวิจัยนี้ได้นำวิธีการวัดแบบโคไซน์มาประยุกต์ใช้ในการวัดความคล้ายคลึงกันของเว็บเซอร์วิสที่ถูกจำแนกหมวดหมู่แล้ว นอกจากนี้ในการจำแนกหมวดหมู่

เว็บเซอร์วิสที่ให้บริการด้านการท่องเที่ยวในอนาคต สามารถนำโมเดลของการจำแนกหมวดหมู่ที่ได้จากเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree) ซึ่งเป็นเทคนิควิธีที่มีประสิทธิภาพมากที่สุดไปใช้ในการทำนายหรือจำแนกหมวดหมู่ของเว็บเซอร์วิสใหม่ที่ให้บริการที่เกี่ยวข้องด้านการท่องเที่ยวได้ และสามารถนำโมเดลนี้ไปเป็นแนวทางในการประยุกต์ใช้เพื่อจำแนกหมวดหมู่ของข้อมูลประเภทอื่น ๆ ได้

5.3 ข้อเสนอแนะ

5.3.1 การปรับปรุงข้อมูลเว็บเซอร์วิสหรือคำศัพท์ในคลังคำในฐานความรู้แบบออนโทโลยีสามารถทำได้โดยอัตโนมัติ เช่น สามารถปรับปรุงข้อมูลได้โดยผ่านทางโปรแกรมเว็บบราวเซอร์ (Web Browser) โดยไม่ต้องกระทำผ่านทางโปรแกรมโพรเทเจ (Protégé) เป็นต้น

5.3.2 การคำนวณหาค่าคุณภาพการบริการของเว็บเซอร์วิสในแต่ละกลุ่มเว็บเซอร์วิสสามารถทำได้โดยอัตโนมัติ โดยผู้ดูแลระบบสามารถปรับตั้งค่าให้อัพเดตข้อมูลได้อย่างอัตโนมัติตามวันเวลาหรือช่วงเวลาที่กำหนดเพื่อให้ค่าคะแนนคุณภาพการบริการของเว็บเซอร์วิสนั้นทันสมัยอยู่เสมอ

5.3.3 ในอนาคตกลุ่มของการท่องเที่ยวอาจมีการปรับเปลี่ยนไปตามสถานการณ์โลก โดยเฉพาะข้อมูลที่ได้จากองค์การการท่องเที่ยวโลกแห่งสหประชาชาติ (United Nations World Tourism Organization : UNWTO) ซึ่งข้อมูลที่มีความน่าเชื่อถือเพราะเป็นองค์กรที่มีบทบาทสำคัญในด้านการท่องเที่ยว ดังนั้นเพื่อให้แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยวนี้มีประสิทธิภาพมากขึ้น ควรปรับเปลี่ยนกลุ่มของการท่องเที่ยวไปตามสถานการณ์การท่องเที่ยวโลก

5.3.4 ควรปรับปรุงส่วนติดต่อกับผู้ใช้ (User Interface) ที่ใช้ในการค้นหาเว็บเซอร์วิสเชิงความหมาย ให้สามารถรองรับการทำงานได้บนอุปกรณ์สมาร์ทโฟน (Smart Phone) และแท็บเล็ต (Tablet) รวมทั้งสามารถใช้งานได้ทั้งบนระบบปฏิบัติการแอนดรอยด์ (Android) และไอโอเอส (iOS) โดยทำเป็นโปรแกรมประยุกต์แบบเคลื่อนที่ (Mobile Application) สามารถค้นหาเพื่อดาวน์โหลดและติดตั้งได้ง่ายผ่านทางเพย์สโตร์ (Play Store) หรือแอพสโตร์ (App Store) เพราะปัจจุบันอุปกรณ์ที่ผู้ใช้ทั่วโลกใช้ในการเข้าถึงอินเทอร์เน็ตเป็นอุปกรณ์เคลื่อนที่มากกว่าคอมพิวเตอร์ส่วนบุคคล (Cullen, www, 2016)

รายการอ้างอิง

กิตติ ภักดีวัฒนะกุล. (2550). **คัมภีร์ระบบสนับสนุนการตัดสินใจและระบบผู้เชี่ยวชาญ**. กรุงเทพฯ: เคทีพีคอมพ์แอนด์คอนซัลท์.

ไกรศักดิ์ เกสร. (2554). **การค้นหาข้อมูลเชิงความหมาย: แนวคิดใหม่ของโปรแกรมค้นหา (Search Engine) และแนวทางการพัฒนาในอนาคต**. วารสารวไลยอลงกรณ์ปริทัศน์ ปีที่ 1 ฉบับที่ 2 กรกฎาคม-ธันวาคม 2554: หน้า 1-16.

นิตยา เกิดประสพ. (2547). **เอกสารประกอบการสอนเรื่องการค้นหาความรู้และการขุดค้นข้อมูล (423304: Knowledge Discovery and Data Mining)**. นครราชสีมา: มหาวิทยาลัยเทคโนโลยี-สุรนารี.

บุญชม ศรีสะอาด. (2543). **การวิจัยทางการวัดผลและประเมินผล**. กรุงเทพฯ: สุวีริยาสาสน์.

พนิดา ตันศิริ (2553). **เว็บเชิงความหมายของเว็บ 3.0**. วารสารนักบริหาร ปีที่ 30 ฉบับที่ 4 หน้า 48-55. กรุงเทพฯ: มหาวิทยาลัยกรุงเทพ.

มารุต บูรณรัช, สภา จรรยาชัชวาล, พรพิมล ผลินกูล, เทพชัย ทรัพย์นิธิ และสุพล ไกลถิ่น. (2553). **โครงการวิจัยระบบสืบค้นข้อมูลแบบบูรณาการเชิงความหมายสำหรับข้อมูลภายในองค์กร**. NECTEC.

วิฑูรย์ ตันศิริคงคล. (2542). **AHP กระบวนการตัดสินใจที่ได้รับความนิยมมากที่สุดในโลก**. กรุงเทพมหานคร: ซีเอ็ดยูเคชั่น.

ศุภกฤษฎิ์ นิวัฒนากูล. (2556). **รายงานวิจัยเรื่อง การเข้าถึงความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย**. สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี.

สมเกียรติ สอนนวล, งามนิจ อาจอินทร์, สมจิตร อาจอินทร์ และ ธีระยุทธ ทองเครือ. (2558). **การสืบค้นเว็บเซอร์วิสแบบสื่อความหมายโดยใช้ซอฟต์แวร์เอเจนต์**. วารสารวิศวกรรมสาร มก. ปีที่ 28 ฉบับที่ 92 หน้า 71-78.

สำนักงานปลัดกระทรวงการท่องเที่ยวและกีฬา. (2559). **รายงานภาวะเศรษฐกิจท่องเที่ยว**, ฉบับที่ 4 เดือนเมษายน-มิถุนายน 2559. โครงการจัดทำรายงานภาวะเศรษฐกิจการท่องเที่ยวรายไตรมาส กระทรวงการท่องเที่ยวและกีฬา.

สุธรรม อุมาแสงทองกุล. (2541). **การศึกษาผลการสืบค้นสารสนเทศที่ใช้เทคนิคดรรชนีแบบกำหนดค่าน้ำหนักที่จัดทำโดยมนุษย์**. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต สาขาวิชาบรรณารักษศาสตร์และสารนิเทศศาสตร์ มหาวิทยาลัยธรรมศาสตร์.

อนุวัฒน์ โยวบุตร และฉัตรตระกูล สมบัติธีระ. (2552). **การค้นหาข้อมูลด้วยการประยุกต์ใช้เว็บเชิงความหมาย**. การประชุมวิชาการระดับชาติด้านวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศและการสื่อสาร ครั้งที่ 3 (The 3rd National CS&ICT Conference: NCSICT 2009)

อรวรรณ อุไรเรืองพันธุ์ และสมจิตร อาจอินทร์. (2552). **การสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยี**. การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 5 (The 5th National Conference on Computing and Information Technology: NCCIT 2009). หน้า 294-299.

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.

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.

Ajao, T. A., Deris, S., and Obasa, I. A. (2013). **QoS-based Web Service Selection Using Filtering, Ranking and Selection Algorithm**. International Journal of Scientific & Engineering Research, Vol. 4, Issue 7, pp.184-190.

Al-Masri, E. and Mahmoud, Q. H. (2007). **QoS-based Discovery and Ranking of Web Services**. IEEE 16th International Conference one Computer Communications and Networks (ICCCN '07), pp.529-534.

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.

Berry, M. J., and Linoff, G. S. (2004). **Data Mining Techniques For Marketing, Sales, and Customer Relationship Management**. Indianapolis: Wiley Publishing, Inc.

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.

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.

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.

Chandrasekaran, B., Josephson, J. R. and Benjamins, V. R. (1999). **What are ontologies, and why do we need them?**. IEEE Intelligent systems, 14(1): pp.20-26.

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.

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.

Cristianini, N. and Shawe-Taylor, J. (2000). **An Introduction to Support Vector Machines and other kernel-based learning methods**, Cambridge University Press.

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

Domingos, P., and Pazzani, M. (1997). **On the Optimality of the Simple Bayesian Classifier under Zero-One Loss**. Machine Learning 29:103-130.

Fielding R.T. (2000). **Representational State Transfer (REST)**, Chapter 5: Architectural Styles and the Design of Network-based Software Architectures. Doctoral Thesis, University of California, Irvine, USA.

Guha, R., McCool., R. and Miller, E. (2003). **Using the Semantic Web: Semantic Search**. Proceedings of the 12th international conference on World Wide Web, New Orleans, 2003, pp.700-709.

Han, J. and Kamber, M. (2001). **Data Mining: Concepts and Techniques**. San Diego: Academic Press.

Han, J., Kamber, M. and Pei, J. (2011). **Data Mining: Concepts and Techniques**. San Francisco: Morgan Kaufmann Publishers.

Hulth, A. (2003a). **Improved automatic keyword extraction given more linguistic knowledge**. Proceedings of the 2003 conference on Empirical methods in natural language processing, pp. 216-223.

Hulth, A. (2003b). **Reducing False Positives by Expert Combination in Automatic Keyword Indexing**. Proceedings of Recent advances in natural language processing (RANLP 2003), pp.367-376.

Kanungo, T., Mount, D. M., Netanyahu, N., Piatko, C., Silverman, R. and Wu, A. (2002). **An Efficient k-Means Clustering Algorithm: Analysis and Implementation**, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, pp.881-892.

Karthiban, R. (2014). **A QoS-Aware Web Service Selection Based on Clustering**. International Journal of Scientific and Research Publications, Vol. 4, Issue 2, pp.1-5.

Kuriakose, J. (2009). **Understanding and Adopting Semantic Web Technology**. Cutter IT Journal. Cutter Information Corp. 22 (9): pp.10–18.

Li, Y., L, Yao., Z, Liangjie., Li, G., Xie, B., and Sun, J. (2007). **An Exploratory Study of Web Services on the Internet**. In IEEE International Conference on Web Services, pp.380-387.

Malczewski, J. (1999). **GIS and Multicriteria Decision Analysis**. New York: John Wiley & Sons.

Manning, C. D., Raghavan, P. and Schütze, H. (2008). **Introduction to Information Retrieval**. London, United Kingdom: Cambridge University Press.

Menasce, D.A. (2002). **QoS issues in Web Services**. IEEE Internet Computing, Vol. 6, No. 6, pp.72-75.

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.

Miller, G. A. (1995). **WordNet: a lexical database for English**. Communications of the ACM, Vol. 38, No. 11, pp.39-41.

Mizoguchi R., Sunagawa E., Kozaki K. and Kitamura Y. (2007). **A Model of Roles within an Ontology Development Tool: Hozo**. Journal of Applied Ontology, Vol. 2, No. 2, pp.159-179.

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.

มหาวิทยาลัยเทคโนโลยีสุรนารี

Muruganadam, M., Gopu, K., and Shanmuga P. R. (2014). **SWSRS: Semantic Web service Retrieval System based on Functional and QOS Evaluation**. International Journal of Engineering Research & Technology (IJERT), Vol. 3 Issue 1, January - 2014, pp.3257-3264.

Musen, M.A. (2015). **The Protégé project: A look back and a look forward**. AI Matters, Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4): pp.4-12.

Naenudorn E. and Niwattanakul S. (2014). **Web Services Selection Algorithm Based on the Quality of Service**. Lecture Notes on Software Engineering, Vol. 2, No. 1, February 2014.

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.

Noy, N. F. and McGuinness, D. L. (2000). **Ontology Development 101: A Guide to Creating Your First Ontology**. California : Stanford University.

Parhi, M., Pattanayak, B. K. and Patra M. R. (2014). **A Multi Agent Based QOS-Driven Web Service Discovery and Composition Framework**. ARPN Journal of Engineering and Applied Sciences, Vol. 9, No. 4, April 2014: pp.542-553.

Perone, C. S. (2013). **Machine Learning :: Cosine Similarity for Vector Space Models (Part III)** [On-line]. Available: http://blog.christianperone.com/2013/09/machine-learning-cosine-similarity-for-vector-space-models-part-iii/

Quinlan, J.R. (1993). **C.45 Programs for machine learning**. San Francisco, CA: Morgan Kaufmann.

Rajendran, T., Balasubramanie, P., and Cherian R. (2010). **An Efficient WS-QoS Broker Based Architecture for Web Services Selection**. International Journal of Computer Applications, Vol. 1, No. 9: pp.79-84.

Russell, J. and Roseby, P. (2006). **Dictionary of Leisure, Travel and Tourism (3rd edition)**. A & C Black Publishers Ltd. United Kingdom.

Salton, G. (1988). **Automatic text processing**. Addison Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Shakhnarovich, G., Darrell, T. and Indyk, P. (2006). **Nearest-neighbor methods in learning and vision : theory and practice**. MIT Press, Cambridge, MA.

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.

SKOS. (2009). **Simple Knowledge Organization System Reference - W3C Recommendation** [On-line]. Available: http://www.w3.org/TR/skos-reference/

Staab, S., Studer, R., Schnurr, H-P. and Sure, Y. (2001). **Knowledge Processes and Ontologies**. IEEE Intelligent Systems. 16(1): pp.26-34.

Sugumaran, V. and Storey, V.C. (2002). **Ontologies for conceptual modeling: their creation, use, and management**. Data & Knowledge Engineering. 42(3): pp.251-271.

Susila, and Vadivel, S. (2014). **QoS Measurement Tool for Web Service Selection**. International Journal of Web Engineering 2014, 3(1): pp.1-8.

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.

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.

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.

UDDI-OASIS Standard. (2004). **Universal Description Discovery and Integration** [On-line]. Available: http://www.uddi.org/

UNWTO. (2015a). **UNWTO Tourism Highlights (2015 Edition)** [On-line]. Available: http://www.e-unwto.org/doi/pdf/10.18111/9789284416899

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.

W3C-Semantic Web Activity. (2013). **World Wide Web Consortium: Semantic Web Activity** [On-line]. Available: http://www.w3.org/2001/sw/

W3C-SOAP. (2007). **World Wide Web Consortium: Simple Object Access Protocol version 1.2 Part 1: (Second Edition)** [On-line]. Available: https://www.w3.org/TR/wsdl20/

W3C-SPARQL. (2013). **World Wide Web Consortium: SPARQL 1.1 Overview** - W3C Recommendation [On-line]. Available: http://www.w3.org/TR/sparql11-overview/

W3C-WS. (2004). **World Wide Web Consortium: Web Services Architecture** [On-line]. Available: https://www.w3.org/TR/ws-arch/

W3C-WSDL. (2007). **Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language** [On-line]. Available: https://www.w3.org/TR/soap12/

W3C-WSMO. (2005). **World Wide Web Consortium: Web service modeling ontology (WSMO)**. W3C Member Submission [On-line]. Available: http://www.w3.org/ Submission/WSMO/

W3C-XML. (2000). **World Wide Web Consortium: Extensible Markup Language (XML) version 1.0 (Second Edition)** [On-line]. Available: http://www.w3.org/XML/

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.

Witten, L. H. and Frank, E. (2005). **Data Mining: Practical Machine Learning Tools and Techniques**. San Francisco: Morgan Kaufmann Publishers.

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.

Zhang, T. (2014). **QoS-aware Web Service Selection based on Particle Swarm Optimization**. Journal of Networks, Vol. 9, No. 3, March 2014: pp.565-570.

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.

ภาคผนวก

แบบสอบถามสำหรับผู้เชี่ยวชาญ

แบบสอบถามสำหรับผู้เชี่ยวชาญ

เรื่อง แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว

แบบสอบถามชุดนี้จัดทำขึ้นเพื่อใช้ประกอบการวิจัยเรื่อง “แบบจำลองการคัดเลือกเว็บเซอร์วิสเชิงความหมายโดยคำนึงถึงคุณภาพการบริการสำหรับการท่องเที่ยว” ขอความกรุณาผู้เชี่ยวชาญโปรดตอบให้ครบทุกข้อ ผู้วิจัยขอขอบพระคุณท่านเป็นอย่างยิ่งที่ให้ความอนุเคราะห์ในการตอบแบบสอบถามครั้งนี้เป็นอย่างดี

แบบสอบถามมีทั้งหมด 76 ข้อ จากข้อมูลชื่อและคำอธิบายเว็บเซอร์วิสต่อไปนี้ ให้ทำเครื่องหมาย ☑ หน้าคำสำคัญที่ท่านคิดว่ามีส่วนเกี่ยวข้องกับเว็บเซอร์วิสนั้น

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ			
1	Airport information	Get Airport Code, CityOrAirport Name, Country, Country Abbrv, CountryCode,GMT Offset Runway Length in Feet, Runway Elevation in Feet, airport code, airport name or ISO country code	☐ hostel	☐ booking	☐ resort	☐ flight
			☐ airway	☐ tour	☐ tourism	☐ latitude
			☐ longitude	☐ celsius	☐ temperature	
2	Jasons	Jasons Travel Media is a major distributor of travel information for New Zealand, Australia, and the South Pacific. Site visitors can explore activities and book accommodation as well as read travel articles, check out destination guides, and share experiences through reviews. Jasons has a Map Helper API that provides developers with programmatic access to information such as the towns in a given area, nearby attractions, and local accommodations.	☐ hostel	☐ booking	☐ resort	☐ flight
			☐ airway	☐ tour	☐ tourism	☐ latitude
			☐ longitude	☐ celsius	☐ temperature	
3	HotelClub	HotelClub is a global accommodation website that offers over 74,000 accommodation options in 141 countries. HotelClub members earn rewards on every booking which can then be redeemed on hotel stays. Members move up in status as they make more bookings, increasing their benefits as they do so.	☐ hostel	☐ booking	☐ resort	☐ flight
			☐ airway	☐ tour	☐ tourism	☐ latitude
			☐ longitude	☐ celsius	☐ temperature	

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
		HotelClub's services are available in 19 currencies and 15 languages. The HotelClub API retrieves information from the site, including a list of all hotels in its database.	
4	Hotelston	The service provides search and reservation booking for travel accommodations. It offers access to hotel room reservations, but also condos, villas, and apartments for short-term rental, depending on the destination. Services include identifying properties by date and location, booking reservations, and canceling previous bookings. API methods support submission of a location with check-in and check-out dates. Returned data include available properties, specifications, and rates with additional functionality to reserve accommodations.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
5	Mapme	Mapme API is a RESTful builder tool for maps. Mounted as a service on Google Maps API, Mapme facilitates easy creation and customization of content on maps. Mapme maps can include layers, tags, moderation for crowdsourcing, activity feed, privacy settings, search, mobile compatibility and more. The API that comes with detailed documentation and full support conveys both HTTPS and HTTP requests, and returns responses in JSON format.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
6	Singapore Metro Rapid Transit	Singapore Metro Rapid Transit is a service that provides information on Singapore's Metro Transit, such as trains, buses, taxis. The SMRT API allows developers to access and integrate the functionality and data of SMRT with other applications and to create new applications. Some example API methods include retrieving bus arrival times, retrieving timetables, and accessing bus service.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
7	Hailo	Hailo is an app that eases the process of hailing a taxi, allowing a user to view and book nearby drivers with a single button. The Hailo API returns nearby driver and ETA information. It also allows the ability to hail a driver from within a 3rd party app, users to 60,000 taxis and executive cars in over 3 continents.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
		Deep linking gives the ability to migrate inputted information seamlessly to the Hailo for a clean user experience.	
8	Kiri	Kiri API provides a programmatic interface to the Kiri Indonesian rideshare platform. Using an API key, developers can implement POST and GET requests over HTTP with JSON serialization to access public transportation options and data. The service is recommended both for local citizens and also for tourists. Users can enter their travel destination, and Kiri automatically suggests a route using angkots (minibuses) as transportation. Kiri's general mission is to help reduce carbon footprints by collaborating in shared transportation efforts.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
9	GeoBulk Reverse Geocoding	GeoBulk offers a simple reverse geocoding API. The API accepts longitude and latitude coordinates over HTTP GET requests, and returns a specific address as a JSON serialized object. The API offers support throughout the U. S. International coverage is not currently supported. The API is free for up to 1000 Reverse Geocodes per day, with increased usage requiring a monthly subscription.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
10	Foursquare Real-time	Foursquare is a local search and discovery service company that provides customers with information on events, restaurants, tourist attractions, and more..The Real-time API allows developers to integrate Real-time information into their applications. There are two sub APIs with this API, one is User Push API which lets the application knows whenever a user has checked in a place.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
11	FlitWays	The FlitWays online booking engine and app allows travelers to pre-book a ride and choose a preferred ride from a selection of vehicles such as Taxi, Executive Cars, Airports Shuttles and Transfers. This gives businesses and private travelers the ability to plan ahead and accordingly when making arrangements for a trip. With FlitWays, travel suppliers such as Airline,Car Rentals, Cruises, Travel Consolidators, Travel Search.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
12	Shellware Weather	The East Coventry Weather Service reports on the current weather conditions in East Coventry, Pennsy.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
13	Weather Web Service	Weather Web Service	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
14	myWeather	Free API for delivery of XML and JSON weather feeds for global locations specified by latitude/longitude, Postcode, or Zipcode. Options include current weather condition and 2-day forecast information for multiple global locations, 7-day forecast information for any single global location, and snow report data plus 6-day forecast information for any single ski resort or mountain location.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
15	NCEP North American Mesoscale	The National Centers for Environmental Prediction (NCEP) North American Mesoscale Model retrieves and analyzes data taken from across the continental U.S. to make predictions about future weather conditions. Data used by this model comes from the Advanced Weather Interactive Processing System . The SOAP API for accessing the NCEP NAM functionality is provided by CUAHSI Hydrologic Information System .	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
16	Dark Sky Forecast	The Dark Sky API integrates weather predictions into applications via JSON & REST protocol. API methods are avialable for Current conditions. Minute-by-minute forecasts out to 1 hour. Hour-by-hour forecasts out to 48 hours and Day-by-day forecasts out to 7 days. The forecasts are available in multiple languages. First 1,000 API calls are available at no cost, following with a fee of $0.0001 for each call after. Dark Sky asks developers to add a “Powered by Forecast” badge that links to http://forecast.io/ where API data is presented.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
17	Weather	Weather forecasting service	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
18	Weather Forecast Service	Weather forecasting service	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
19	Lost Springs Weather	Local weather information service	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
20	World Weather Online	The Premium Local Weather REST API method allows you to access current weather conditions and the next 15 days of accurate and reliable weather forecast. The Local Weather API returns weather elements such as temperature, precipitation (rainfall), weather description, weather icon and wind speed.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
21	GlobalWeather	Current weather and weather conditions for major cities around the world	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
22	DOTS Fast Weather	Weather forecasting service	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

มหาวิทยาลัยเทคโนโลยีสุรนารี

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
23	OpenWeather	Access current weather data for any location on Earth including over 200,000 cities! Current weather is frequently updated based on global models and data from more than 40,000 weather stations. Data is available in JSON, XML, or HTML format.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
24	US Weather	Get five day weather report for a given zipcode	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
25	Baron Velocity Weather	The Velocity Weather API allows developers to incorporate the Baron Weather service into their applications. Data products in the API fall into two categories – Text and Graphical products. Text products are returned in either JSON text formats, and can be displayed on a map or integrated into a user interface. Graphical products are images, provided in PNG format for overlay onto a map. Data provided includes shore wave conditions, tornado tracking, roadway conditions, Baron provides actionable and intelligent weather services helping business across several industries mitigate dangerous weather.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
26	WeatherForecast	Get one week weather forecast for valid zip code or Place name in USA	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
27	Weather Forecasting Service	Weather forecasting service	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
28	Yahoo Weather	The Yahoo Weather RSS feed enables you to get up-to-date weather information for your location. This data can be used as an RSS feed in a feed reader or can be integrated into third party applications. The Weather RSS feed is a dynamically-generated feed based on zip code or Location ID.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
29	Five Day Weather	The 5 Day Weather API accepts a city as a single parameter and returns a 5 day forecast. Specifically, the API will accept an HTTP request with a city name identifier and will return a JSON object that includes the requested city name, the temperature, the sky conditions, projected humidity, wind speed in MPH, and day names with associated calendar dates. Users can make specific calls to switch between Celsius and Fahrenheit and Km/h and mph according to local standards. The API is free to use without restriction.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
30	FareBookings Airport	Farebookings provides a set of tools to help travels and travel companies. The tools include a taxi fare calculator, a booking service, and other tools. Farebookings exposes it's data on airports through a REST API. The FareBookings Airport Data API has been developed to provide usefull Airports information.The APi is provided for free. Users must provide a IATA AIRPORT CODE, and you will receive information about this Airport, name in different languages and other information.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
31	British Airways	British Airways (BA) is a full service global airline with an extensive global route network based out of the United Kingdom. British Airways offers developers a REST API for the integration of BA data into third party applications. Current public API offerings include: Flight Time Info, for the latest flight information distributed by BA; Flight Offer, for the latest BA flight offers; and Availability, for hotel, car hire, and package offer information distributed by BA.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
32	Urgent	The Urgent.ly API allows developers to access and integrate the functionality of Urgent.ly with other applications and to create new applications. Public documentation is not available; interested developers should contact Urgent.ly for more transportation. Urgent.ly is an on-demand roadside assistance application.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
33	Cabforce	Cabforce allows users to search for and book taxi rides and airport transfers online. They provide services for all major European travel destinations and are currently expanding into the U.S., Asia, and Australia. Cabforce works with certified service providers who operate clean, non-smoking cars and are guaranteed to arrive on-time. Bookings are always con⊙ฉrmed, and Cabforce provides 24/7 customer service in English.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
34	Waynaut	Waynaut API provides a RESTful JSON interface to implement mapping features around Europe. With this service, developers have access to door to door search and multimodal trip planning integration. Waynaut allows users to find the best way to their destination using combined traditional methods of transportation such as bus, taxi with innovative methods such as ridesharing. To see how Waynaut works with trains, cars, BlaBla cars.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
35	Transport API	Transport API is a data-as-a-service company that provides transportation data to apps and websites using industry sources like Travelline, Network Rail. The Transport API uses REST calls and has a variety of endpoints. Users can make calls to get data on bus routes and times, London underground transit info, several public trip planners. The API is already being used for live bus departure services, UK travel option applications.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

มหาวิทยาลัยเทคโนโลยีสุรนารี

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
36	Taxi Fare Finder	The value of Taxi Fare Finder API is to find the rate between two locations. As the fee might vary according to every place, the site features a list of more than 20 entities that include United States, UK, Australia, Netherlands, Belgium, New Zealand, India, Spain and China. This API has a secure connection via SSL. Some of the properties in JSON protocol cover name, locale and distance. To get a taxi fare, the required parameters are key, origin and destination, while the optional parameters are callback and traffic.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
37	Airport Transfer Worldwide	Airport Taxi Transfer Worldwide is a web based international transfer service, which offer transfers to and from airports, railway stations and more. The service allows users to avoid long waits at airports or having to travel on public transit. Transfers are offered at fixed prices and are door-to-door in nature.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
38	Amadeus Rail Station	The Amadeus Rail-Station Information API provides developers a way to integrate the Amadeus Rail-Station Information service with their applications, enabling their users to retrieve the rail station information corresponding to an Amadeus UIC rail station ID. Only French and Italian stations are supported currently. Amadeus is a technology company that provide services businesses and individuals relating to their travels.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
39	MyAgentDeals	MyAgentDeals is a global travel product provider that has sourced a broad portfolio of Airfare, hotel accommodations, transfer services, city tours, tour guides, excursions, and over 90,000 hotel properties in 145 countries. The MyAgentDeals XML API allows travel vendors to access the MyAgentDeals database of travel content for reservation systems or third-party websites and applications. The service uses REST calls and returns in XML formats.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

มหาวิทยาลัยเทคโนโลยีสุรนารี

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
40	GTA Travel	GTA-Travel is a global travel product provider that has sourced a broad portfolio of accommodation options, transfer services, city tours, tour guides, excursions, restaurants, and over 45,000 vendors in 185 countries. The GTA XML API allows travel vendors to access the GTA database of travel content for reservation systems or third-party websites and applications. The services uses REST calls and returns XML. An account is required with service.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
41	Amusement Markup Language	Amusement Markup Language API gathers information on attractions, events and points of interest around the world. It's goal is to allow for searching by keyword or zip code to find different points of interest and attractions within an amusement park.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
42	to.uri.st	Use the to.uri.st API to access tourist information data from to.uri.st. to.uri.st is about finding those local attractions you never knew existed, we spider the Web for places to go the World over and place them on our map for you to discover. They have a Web spider that fetches data from different tourist information sites around the World. It is an Atom-based HTTP API. All to.uri.st data is free under the Creative Commons Attribution-Share Alike 3.0 License.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
43	Tourism New Zealand	Tourism New Zealand website contains inspiring, up-to-date travel information about destinations, attractions and activities in New Zealand. The API makes the site's data available to users.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
44	Google Places	The Google Places Web Service is a service that returns information about a place: an establishment or prominent point of interest using an HTTP request. Two basic Place requests are available: Place Search request and a Place Details request. Generally, a Place Search request is used to return candidate matches, while a Place Details request returns more specific information about a attraction.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
45	Trippin	Trippin'in offers hourly changing (!) restaurant recommendations, nightlife recommendations, outdoors and attraction recommendations via API. They disrupt the Travel and local market by offering a new method for local recommendations, replacing the known "review-based" method offered by companies like Tripadvisor and Yelp. By measuring social interaction (such as Likes, Check-ins, and Tweets) between people and places (like restaurants, nightclubs, or bars) their algorithm can determine how popular (or crowded) those places will be, for any hour of any day.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
46	ValueCommerce	ValueCommerce is a Japanese Affiliate Marketing and Search Marketing Agency Service. It offers the Travel API which allows users to aggregate airfare, hotel and tour from top merchants in Japan. The API allows users the following functionality: do a quick Search by area and price of hotel accommodations, find more information on hotel accommodation and room types.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
47	Myallocator	Myallocator.com is a platform for hotels, hostels, and other lodging companies to update their accommodations and information on many online travel sites. The Myallocator.com API allows developers to access and integrate the functionality of Myallocator.com with other applications and to create new applications. Some example API methods include managing listings, retrieving price and accommodation information, and managing users.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
48	Zumata Hotels	Zumata is a web service providing ancillary revenue solutions for hotels, airlines, travel agents, event managers, tourism boards, and other travel industry entities. The Zumata Hotels API provides programmatic access to Zumatas integratable booking engine, allowing customer applications to find, book, and cancel hotel reservations.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
49	Checkfront	Checkfront is a booking service that allows businesses to manage their schedule based inventories, centralize reservations, and process payments online. It can be used to manage accommodation, events, rentals and professional service bookings.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
50	One Map Singapore	The One Map Singapore API allows to integrate mapping features into applications. It provides REST protocols and Api Key authentication. One Map invites developers to create mashups with the JavaScript APIs available on the official site, that include Address Search, Agency Data Search, Static Map, Mashup, and Advanced Mashup.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
51	Yelp	The Yelp APIs are RESTful APIs and users can retrieve business review and rating, information for a particular geographic region or location.display review information for a particular business, determine accurate neighborhood name information for a particular location, track recent reviews for a particular business, display pictures of highly rated local businesses and of the top reviewers for that business, determine a particular business' review and rating information based on the phone number for that business. The default output is JSON. This output format was chosen due to the availability of JSON parsers in many languages.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
52	Rezgo	Rezgo is tour operator software designed for companies who run tours, activities, and attractions. Many of the companies using Rezgo include whale watching tours, river rafting, kayaking tours, bicycle tours, city guided tours, bus tours, boat tours, helicopter tours, special events, and more. Rezgo helps tour operators organize their tours and gives them all the tools. As a tour operator system, Rezgo helps tour operators increase their online while keeping them organized and reducing their administrative burden and costs. Rezgo supports inventory	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
		management, realtime availability, rich media like photos videos, pricing and availability rules and online vouchering.	
53	Travel Booking Network	Travel Booking Network is an accommodation and ticketing platform for bed and breakfasts, hotels, hostels, university residences, guest houses, and other holiday rentals. They process bookings through their own websites, but also lend their technology out to other accommodation providers and travel websites. Travel Booking Network can set up either an API or a white label service for a client's website.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
54	Close To The Point	Close To The Point (CTTP) provides geolocation data so that app providers can better interact with their user location and demographics. The CTTP REST API will provide real time statistics on point of interests so that providers can have a comprehensive monitoring solution for their app's usage. Location data comes in the form of real time and historic so providers may witness how often users frequent certain locations. In addition to the REST API implementation, developers can add the CTTP Android & iOS libraries to their apps.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
55	Ola Ride Availability	The Ola Ride Availability API allows developers to access information on ride availability by providing the user's latitude and longitude location in India. By making calls to the API, users can access this information directly from the application.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
56	Expedia	Expedia is a world leader in travel and technology, and the EAN Developer Hub gives developers free access to a set of APIs that power websites, mobile apps, and much more. The EAN APIs can be used to book hotels, flights and cars. The APIs use both REST and SOAP protocol and responses are formatted in XML	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
57	BeaconsInSpace	The BeaconsInSpace REST API provides developers with access to beacons across establishments in major cities for integration with other applications. Some example API methods include retrieving lists of beacons UUIDs, retrieving information about beacons, and retrieving beacons by geographic coordinates. BeaconsInSpace provides information and access to deployed beacons to build Internet of Things and location-based applications.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
58	Roomorama	Roomorama is a service for travellers to find short-term accommodations in locations around the world. Property owners can list their properties/rooms for short-term rental that travellers can search for and make reservations. The Roomarma API allows developers to access the data and functionality of Roomarama. Some example API methods include listing destinations, properties, retrieving, creating, and editing property information, and listing property availability and rates.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
59	Cleartrip Hotel	Cleartrip provides booking for flights, hotels and trains, as well as other travel services across the world. The Cleartrip Hotel API provides developers access to the hotel search and booking functionalities of Cleartrip. You can search for hotels, get more information, book a hotel, view itenerary and even cancel hotels, all from the Cleartrip Hotel API interface.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
60	MapFruition	The MapFruition Geocoder API lets developers transform addresses and place names into geographic coordinates. The reverse geocoder enables turning a geographic coordinate pair into a meaningful address and/or place name. All valid results are returned in GeoJSON format. Data coverage currently includes the Contiguous US, Alaska, Hawaii, Mexico, and Canada. MapFruition is a provider of geospatial, GIS/GPS, and technical business solutions.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
61	MapBox	Mapbox is a suite of open source mapping tools. This API can be used to create interactive maps using a wide range of file formats and databases. It also allows users to host their maps so that they can be embedded on webpages and shared with a wider audience. All maps created using Mapbox are compatible with Android and iOS devices.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
62	Google Maps	The Google Maps API allow for the embedding of Google Maps onto web pages of outside developers, using a simple JavaScript interface or a Flash interface. It is designed to work on both mobile devices as well as traditional desktop browser applications. The API includes language localization for over 50 languages, region localization and geocoding, and has mechanisms for enterprise developers who want to utilize the Google Maps API within an intranet.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
63	Australian PostCode	Australian Postcode ,Location Web service Get Australian Postcode by partial / full Location Name Get Australian Location by partial / full Postcode Get Australian Postcode by partial / full Postcode Australian Postcode ,Location Web service	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
64	UK Location	Get UK Postcode,Town,Country and Validate UK Address Get UK town, Postcode and County by full /partial County Get UK town, Postcode and County by full /partial Town Get UK town, Postcode and County by Postcode ,Use First Section of Poscode for Postcode atribute Get UK town location	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
65	IP2Geo	Our system uses a multiple sources for resolution. Including a host database, trace route information, and other systems. Resolves IP addresses to Organization, Country, City, and State/Province, Latitude, Longitude. In most US cities it will also provide some extra information such as Area Code, and more.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
66	Toprural	Toprural is a rural accommodation search engine built to promote rural tourism throughout Europe. The Toprural API provides users with GET requests for accommodation listings information, including the ability to search within specified geographical parameters. Responses may be XML or JSON formatted.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
67	Majestic International Group	The Majestic International Group bills itself as a wholesale tour operator for the U.S., Canada, Mexico, the Caribbean, Europe, and Asia. The company specializes in trips to Disney destinations, Sea World parks, and Universal Orlando. It provides information and online booking services for hotels, tours, car rentals, etc. The Majestic International Group API enables users to book hotels as well as search for and retrieve information from the website.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
68	Adriagate	Adriagate is a travel service specializing in holiday accommodations for Croatia. Their online offerings represent a large proportion of the tourism opportunities available. These include private accommodations, hotels, Adriatic cruises, and authentic Croatian accommodations, such as stone houses and remote Robinson cottages on the islands. Additionally, Adriagate offers transportation services that include ferry tickets, transfers, and rent-a-car services.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
69	EasyToBook	Easytobook will release a REST API. Easytobook provides hotel bookings for more than 120,000 hotels in over 9,000 cities. The Easytobook provides a SOAP API to facilitate hotel searching and booking applications. Developers can retrieve a list of cities, get hotels for a particular city, check hotel availability, look for cheap hotel rooms, get other room information, retrieve hotel photos and other hotel information, as well as book the hotel on behalf of their users.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

มหาวิทยาลัยเทคโนโลยีสุรนารี

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
70	Resort Data Processing	The service provides vacation property management functions for hotels, resorts, timeshares and condo rentals, RV resorts, and campgrounds. Services include website hosting, management of accommodation options, reservations, and booking. It also provides tools for backoffice functions like general ledger and auditing. API methods support listing accommodations available with pricing and date selection, with booking of reservations and payment processing. Methods also support reservation lookup, updating, and cancellation.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
71	Global Matrix	The service provides outsourced support for travel resellers, including research on rates and fares, selection of reservation options, and booking of reservations. It aggregates offering information for hotel rooms and more from major GDS providers and presents options detailing possible choices. Applications can format for user selection, then return that data to the service, which processes the reservation.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
72	i-Rent	The service provides functions for vacation property rentals, either by property owners themselves or by property management agencies. Functions include web content management and site hosting, information about specific accommodations, reservations, and booking. Various nationally branded services provide property management in European and North American countries. API methods support creating and updating an accommodation listing on the service, including description.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
73	TransHotel	The service provides hotel search and look-up function with ability to book accommodations. It is designed for integration with travel services to allow selection of accommodations, with reservation booking, cancellation, transfer capabilities. It includes flexible payment options, including direct payment to the properties booked. API methods support submission of location and check-	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

ที่	ชื่อเว็บเซอร์วิส	คำอธิบายเว็บเซอร์วิส	คำสำคัญ
		in/check-out dates with filters for hotel rating, specific amenities(swimming pool, restaurant), and local features.	
74	Alpharooms Affiliate	The Alpharooms Affiliate website helps travelers find discount hotels, budget flights, and cheap holidays worldwide. The site provides access to information on over 250,000 hotels and hundreds of airlines along with car hire and worldwide insurance coverage. The Alpharooms Affiliate API lets users search for available hotel rooms by destination, currency, locale, and several combinations thereof using SOAP calls.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
75	Serko Online	Serko Online is Asia-Pacific's first and only integrated online corporate booking engine, expense management system, and real-time financial reporting tool. It allows corporate customers to handle their own travel bookings and covers the full cycle of travel-based transactions including flights, accommodations, rental car bookings, and insurance.	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature
76	eContent.Maps	eContent.Maps provides an interactive map of housing, events, dining places, tours, and attractions that covers Germany along with parts of Austria and the surrounding countries. Users can set custom markers, find routes, and much more using the web interface or the SOAP API	☐ hostel ☐ booking ☐ resort ☐ flight ☐ airway ☐ tour ☐ tourism ☐ latitude ☐ longitude ☐ celsius ☐ temperature

มหาวิทยาลัยเทคโนโลยีสุรนารี

ประวัติผู้เขียน

นายเอกชัย แน่นอุดร เกิดเมื่อวันที่ 19 ตุลาคม พ.ศ. 2524 ภูมิลำเนาอยู่จังหวัดร้อยเอ็ด สำเร็จการศึกษาระดับปริญญาตรีเมื่อปี พ.ศ. 2546 ในสาขาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม สำเร็จการศึกษาระดับปริญญาโทเมื่อปี พ.ศ. 2549 ในสาขาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น และในปี พ.ศ. 2554 ได้เข้าศึกษาต่อในระดับปริญญาเอก สาขาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี โดยปี พ.ศ. 2554-2557 ได้รับทุนสนับสนุนการศึกษาต่อในระดับปริญญาเอกจากกองทุนพัฒนาบุคลากรมหาวิทยาลัยมหาสารคาม และปี พ.ศ. 2558 ได้รับทุนสนับสนุนจากคณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม

ปัจจุบันเป็นอาจารย์ประจำสาขาวิชาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม

มหาวิทยาลัยเทคโนโลยีสุรนารี