

รหัสโครงการ SUT2-204-54-24-04



รายงานการวิจัย

การออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่าน
ระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย

(The Design and Development of a Knowledge Extraction
and Retrieval System via Online GIS for
Thailand Tourism Business)

ได้รับทุนอุดหนุนการวิจัยจาก
มหาวิทยาลัยเทคโนโลยีสุรนารี

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว



รายงานการวิจัย

การออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่าน
ระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย
(The Design and Development of a Knowledge Extraction
and Retrieval System via Online GIS for
Thailand Tourism Business)

คณะผู้วิจัย

หัวหน้าโครงการ

อาจารย์ ดร. จิตินนต์ อังสกุล
สาขาวิชาเทคโนโลยีสารสนเทศ
สำนักวิชาเทคโนโลยีสังคม
มหาวิทยาลัยเทคโนโลยีสุรนารี

ผู้ร่วมวิจัย

อาจารย์ ดร. ธรา อังสกุล

ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ปีงบประมาณ พ.ศ. 2554-2555

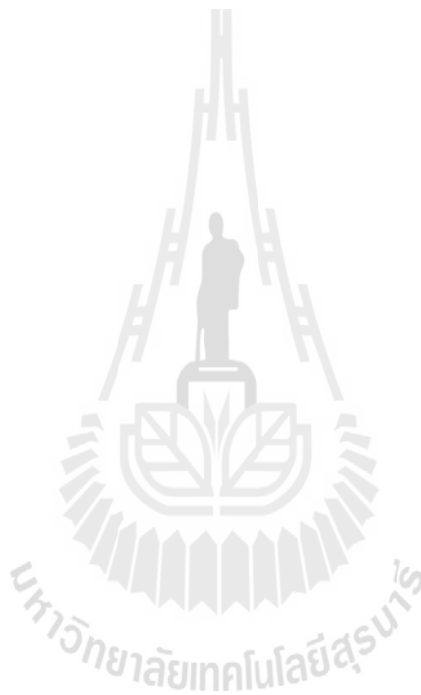
ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

กันยายน 2555

กิตติกรรมประกาศ

ผู้วิจัยต้องขอขอบคุณเป็นอย่างยิ่งต่อมหาวิทยาลัยเทคโนโลยีสุรนารีที่ได้สนับสนุนเงินอุดหนุน การวิจัยประจำปีงบประมาณ พ.ศ. 2554 – พ.ศ. 2555 ให้กับโครงการวิจัยนี้ ซึ่งเป็นส่วนสนับสนุนที่สำคัญยิ่งในการการสร้งองค์ความรู้ และการพัฒนาผลงานวิจัย เพื่อนำไปใช้ให้เกิดประโยชน์ต่อสังคม และประเทศชาติต่อไป

จิตติมนต์ อั่งสกุล



บทคัดย่อ

ปัจจุบันนักทอ่งเที่ยวส่วนใหญ่จะใช้อินเทอร์เน็ตในการสืบค้นข้อมูลเพื่อนำมาใช้ประกอบการตัดสินใจเลือกแหล่งท่องเที่ยวให้ได้ตรงกับความต้องการ ซึ่งวิธีที่นิยมใช้กันโดยทั่วไปคือ การตัดสินใจโดยใช้ข้อมูลจากบทวิจารณ์ของนักทอ่งเที่ยวที่เคยไปมาแล้ว แต่อย่างไรก็ตาม การเลือกแหล่งท่องเที่ยวให้ได้ตรงกับความต้องการนั้น นักทอ่งเที่ยวจะต้องอ่านบทวิจารณ์ทั้งหมดซึ่งมีเป็นจำนวนมาก ในการวิจัยนี้ได้ศึกษาและวิเคราะห์ขั้นตอนวิธีที่เกี่ยวข้องกับการสกัดความรู้แบบอัตโนมัติ ซึ่งเป็นการผสมผสานระหว่างระบบเว็บเชิงความหมายและระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับบริการทอ่งเที่ยว ซึ่งระบบดังกล่าวจะช่วยนักทอ่งเที่ยวในการวางแผนการท่องเที่ยวโดยเฉพาะในเรื่องของที่พักให้เหมาะสมกับความต้องการของนักทอ่งเที่ยวมากที่สุด เนื่องจากได้รับคำแนะนำจากนักทอ่งเที่ยวที่เคยไปมาแล้วผ่านระบบสกัดความรู้ และได้รับสารสนเทศทางภูมิศาสตร์ออนไลน์ซึ่งแสดงสิ่งอำนวยความสะดวกต่าง ๆ รอบที่พักนั้น ๆ นอกจากนี้นักทอ่งเที่ยวยังสามารถแสดงความคิดเห็นเกี่ยวกับที่พักต่าง ๆ เพื่อเป็นประโยชน์ต่อนักทอ่งเที่ยวกลุ่มอื่นได้อีกด้วย

การวิจัยนี้ได้นำเอาเทคนิคต่าง ๆ ที่มีอยู่ในปัจจุบัน เช่น ออนโทโลยี มาใช้เป็นฐานความรู้ในการสกัดและจัดเก็บความรู้ และตัวแปลภาษา มาช่วยในการวิเคราะห์เชิงความหมาย รวมทั้งพัฒนาวิธีการคำนวณระดับคะแนนความพึงพอใจด้วยเทคนิคตรรกศาสตร์คลุมเครือ เพื่อให้ได้มาซึ่งผลลัพธ์ที่มีความถูกต้องมากที่สุด ในการแสดงผลความรู้ที่สกัดได้ ระบบแสดงให้เห็นถึงความสัมพันธ์ของข้อมูลในรูปแบบของแผนภาพโครงสร้างต้นไม้ พร้อมทั้งแสดงจำนวนความคิดเห็นและระดับคะแนนความพึงพอใจของแต่ละคำวิจารณ์อย่างละเอียด นอกจากนี้ระบบยังสามารถให้ผู้ใช้เลือกข้อมูลเพิ่มเติมของสถานที่ต่าง ๆ ของบทวิจารณ์ออนไลน์นั้น ในรูปแบบของระบบสารสนเทศทางภูมิศาสตร์ที่ผู้ใช้สามารถเข้าใจได้ง่ายและนำไปใช้วางแผนได้จริง

ผลการทดลองพบว่า ระบบที่นำเสนอมีความถูกต้องของการสกัดความรู้ร้อยละ 90.15 ค่าความแม่นยำร้อยละ 100 ค่าความระลึกร้อยละ 75.12 ค่าระดับคะแนนของคำวิจารณ์โดยรวมแตกต่างจากระดับคะแนนที่ผู้กำหนดเท่ากับ 0.37 คะแนน และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองมีค่าเท่ากับ 0.466 ทั้งนี้ผู้ประกอบการธุรกิจอิเล็กทรอนิกส์สามารถนำระบบที่พัฒนาขึ้นไปประยุกต์ใช้ในสกัดและค้นคืนความรู้เกี่ยวกับลูกค้า แล้วนำความรู้เหล่านั้นมาพัฒนาสินค้าหรือบริการของตนเองเพื่อตอบสนองความต้องการของลูกค้าให้ได้มากที่สุด

ABSTRACT

Currently, most tourists use the internet to retrieve information for supporting their decision in selecting the tourist places that conform to their preferences. The most common method is the decision based on reviews of experienced tourists. However, tourists must read enormous reviews in order to select their preferred tourist places. This research studied and analyzed the process of an automatic knowledge extraction, which is the integration between semantic web and online geographical information system (GIS) for tourism domain. The developed system could help tourists plan their trips appropriately, especially the accommodation, based on recommendation from experienced tourists via the extraction system. The tourists also obtain online GIS presented with facilities surrounding the accommodation. Moreover, tourists can express their opinion of the accommodations, which could be useful for other tourists in the future.

This study combines several existing techniques, such as using ontology for the knowledge base for knowledge extraction and storage, using parser to solve the semantic analysis problem. In addition, this study proposed a calculation approach for a user satisfaction level based on fuzzy logic in order to achieve maximize accuracy of results. In order to present extracted knowledge, this system illustrates the relationship of data in forms of hierarchy diagram, the number of comments and the user satisfaction level of each review thoroughly. In addition, the system can present additional information of the places in form of online GIS, which is easy to understand and can be used in practice.

The experimental results indicated that the proposed system achieved with 90.15% of accuracy, 100% of precision, 75.12% of recall, 0.37 review points that are different from the point specified by user and 0.466 of root mean square error (*RMSE*). Thus, e-business operators can apply this system to extract and retrieve about knowledge of the customer and use these data to develop their product or service for satisfying the needs of customers as much as possible.

สารบัญ

	หน้า
กิตติกรรมประกาศ.....	ก
บทคัดย่อภาษาไทย.....	ข
บทคัดย่อภาษาอังกฤษ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญภาพ.....	ฉ
บทที่	
1 บทนำ	
1.1 ความสำคัญและที่มาของปัญหาการวิจัย.....	1
1.2 วัตถุประสงค์ของการวิจัย	3
1.3 สมมติฐานการวิจัย.....	3
1.4 ข้อตกลงเบื้องต้น	4
1.5 ขอบเขตของการวิจัย	4
1.6 ประโยชน์ที่คาดว่าจะได้รับ	4
1.7 คำอธิบายศัพท์	5
2 ปรัชญาทฤษฎีและงานวิจัยที่เกี่ยวข้อง	
2.1 การสกัดและค้นคืนความรู้แบบอัตโนมัติ.....	7
2.1.1 ความเป็นมาของการสกัดและค้นคืนความรู้แบบอัตโนมัติ	7
2.1.2 ขั้นตอนวิธีการสกัดและค้นคืนความรู้แบบอัตโนมัติ.....	8
2.2 ทฤษฎีตัวแปลภาษา.....	9
2.2.1 ที่มาและกระบวนการแปลภาษา.....	9
2.2.2 การประมวลผลหน่วยคำ	11
2.2.3 การประมวลผลทางไวยากรณ์.....	12
2.2.4 การประมวลผลทางความหมาย.....	14
2.3 ทฤษฎีออนโทโลยี.....	14
2.3.1 นิยามของออนโทโลยี.....	15
2.3.2 ประเภทของออนโทโลยี.....	16
2.3.3 กระบวนการพัฒนาออนโทโลยี	16
2.3.4 ภาษาที่ใช้ในการพัฒนาออนโทโลยี.....	18
2.3.5 งานวิจัยที่นำออนโทโลยีเข้ามาช่วยในการสกัดและค้นคืนความรู้	18

สารบัญ (ต่อ)

	หน้า
2.4 ระบบสารสนเทศทางภูมิศาสตร์ออนไลน์.....	20
2.4.1 ความหมายของระบบสารสนเทศทางภูมิศาสตร์ออนไลน์.....	20
2.4.2 องค์ประกอบของระบบสารสนเทศทางภูมิศาสตร์ออนไลน์.....	21
2.4.3 กูเกิ้ลแมพเอพีไอ.....	22
2.5 แนวคิดเกี่ยวกับความพึงพอใจของนักท่องเที่ยว.....	22
2.5.1 ความหมายและความสำคัญของความพึงพอใจของนักท่องเที่ยว.....	22
2.5.2 งานวิจัยที่เกี่ยวข้องกับความพึงพอใจของนักท่องเที่ยว.....	24
2.6 แนวคิดเกี่ยวกับบทวิจารณ์ออนไลน์ของนักท่องเที่ยว.....	25
2.6.1 ความสำคัญของบทวิจารณ์ออนไลน์.....	26
2.6.2 งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์บทวิจารณ์ออนไลน์.....	27
3 วิธีดำเนินการวิจัย	
3.1 วิธีวิจัย.....	30
3.1.1 ศึกษาและวิเคราะห์ปัญหาที่เกิดขึ้นในการสกัดและค้นคืนความรู้ ที่มีอยู่ในปัจจุบัน.....	30
3.1.2 ออกแบบและพัฒนาฐานความรู้.....	31
3.1.3 ออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทาง ภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย.....	37
3.1.4 ทดสอบและประเมินผลการทำงานของระบบ.....	53
3.2 เครื่องมือที่ใช้ในการวิจัย.....	53
3.2.1 เครื่องมือที่ใช้ในการพัฒนาระบบ.....	53
3.2.2 เครื่องมือในการพัฒนาออนไลน์.....	54
3.3 การเก็บรวบรวมข้อมูล.....	55
3.4 การวิเคราะห์ข้อมูล.....	55
3.4.1 การวิเคราะห์ผลการประเมินกระบวนการสกัดคุณลักษณะที่สำคัญ จากบทวิจารณ์.....	56
3.4.2 การวิเคราะห์ผลกระบวนการคำนวณระดับคะแนนของที่พัก.....	57
4 ผลการวิจัยและการอภิปรายผล	
4.1 ผลการพัฒนาระบบ.....	58
4.2 ผลการประเมินความถูกต้องของส่วนอนุมานความรู้.....	69

สารบัญ (ต่อ)

	หน้า
4.2.1 ผลการประเมินความถูกต้องของการสกัดคุณลักษณะที่สำคัญของสถานที่พัก	71
4.2.2 ผลการหาค่าระดับคะแนนความสัมพันธ์ระหว่างคำวิเศษณ์และคำคุณศัพท์	77
4.2.3 ผลการประเมินความถูกต้องของการคำนวณระดับคะแนนของสถานที่พัก	77
4.3 การอภิปรายผล	79
5 สรุปและข้อเสนอแนะ	
5.1 สรุปผลการวิจัย	81
5.2 ข้อจำกัดของการวิจัย	83
5.3 การประยุกต์ผลการวิจัย	83
5.4 ข้อเสนอแนะในการวิจัยครั้งต่อไป	83
บรรณานุกรม	84
ภาคผนวก	
ภาคผนวก ก ตัวอย่างข้อมูลที่ได้จากระบบ	91
ภาคผนวก ข การใช้งานโปรแกรม Protégé	148
ประวัติผู้วิจัย	156

สารบัญตาราง

ตารางที่	หน้า
3.1 ปัจจัยที่ใช้ในระบบสกัดและคั่นคั้นความรู้แบบอัตโนมัติและผลลัพธ์ที่ได้.....	31
3.2 ตัวอย่างการให้คะแนนของนักท่องเที่ยวยุคใหม่ในเรื่องทำเลที่ตั้งจากกลุ่มตัวอย่างบทวิจารณ์.....	47
3.3 เกณฑ์การพิจารณาความใกล้เคียงระหว่างระดับคะแนนในเรื่องต่าง ๆ กับระดับ คะแนนในภาพรวมของทั้งโรงแรม.....	47
3.4 เกณฑ์การพิจารณาค่าน้ำหนักจากร้อยละความน่าจะเป็นที่การวิจารณ์เฉพาะเรื่อง หนึ่ง ๆ จะส่งผลกระทบต่อระดับคะแนนในภาพรวม.....	48
3.5 ตัวอย่างกฎตรรกศาสตร์คลุมเครือที่ใช้ในการคำนวณระดับคะแนนในภาพรวม.....	49
3.6 ตัวอย่างการแสดงผลระดับคะแนนเป็นรูปดาว เมื่อจำนวนเต็มของระดับคะแนนทั้ง โรงแรมเท่ากับ 4 คะแนน.....	52
3.7 ข้อมูล 4 ประเภทที่แตกต่างกันตามความเกี่ยวข้องและการสกัดได้.....	56
4.1 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 1.....	71
4.2 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 2.....	72
4.3 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 3.....	72
4.4 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 4.....	73
4.5 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 5.....	74
4.6 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจโดยรวม.....	74
4.7 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) และค่ารากที่สองของ ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ระหว่างวิธีการซึ่งใช้ตรรกศาสตร์ คลุมเครือและวิธีการซึ่งไม่ได้ใช้ตรรกศาสตร์คลุมเครือ.....	78
4.8 การเปรียบเทียบร้อยละของค่าผิดพลาดระหว่างวิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือและ วิธีการซึ่งไม่ได้ใช้ตรรกศาสตร์คลุมเครือ.....	79
ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations).....	92
ก.2 คุณสมบัติทั้งหมดของคลาสห้องน้ำ (Bathroom).....	99
ก.3 คุณสมบัติทั้งหมดของคลาสห้อง (Room).....	100
ก.4 คุณสมบัติทั้งหมดของคลาสห้องพักรับรอง (Guestroom).....	101
ก.5 คุณสมบัติทั้งหมดของคลาสห้องประชุม (ConferenceRoom).....	104
ก.6 คุณสมบัติทั้งหมดของคลาสกิจกรรม (Activity).....	105
ก.7 คุณสมบัติทั้งหมดของคลาสนันทนาการที่อำนวยความสะดวก (Amenity).....	106

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ก.8 คุณสมบัติทั้งหมดของคลาสสถานที่ท่องเที่ยว (Attraction).....	108
ก.9 คุณสมบัติทั้งหมดของคลาสข้อมูลการติดต่อ (ContactData)	110
ก.10 คุณสมบัติทั้งหมดของคลาสวัน เวลา เปิดทำการ (DateTime).....	110
ก.11 คุณสมบัติทั้งหมดของคลาสกิจกรรมพิเศษ (Event)	111
ก.12 คุณสมบัติทั้งหมดของคลาสพิกัด (GPSCoordinates).....	113
ก.13 คุณสมบัติทั้งหมดของคลาสที่อยู่ (PostalAddress)	113
ก.14 คุณสมบัติทั้งหมดของคลาสการจัดอันดับ (Rating).....	113
ก.15 คุณสมบัติทั้งหมดของคลาสตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว (Ticket).....	114
ก.16 คลังคำเหมือนของระบบ	115
ก.17 คำอธิบายป้ายระบุที่ใช้ในคลังคำศัพท์	119

สารบัญภาพ

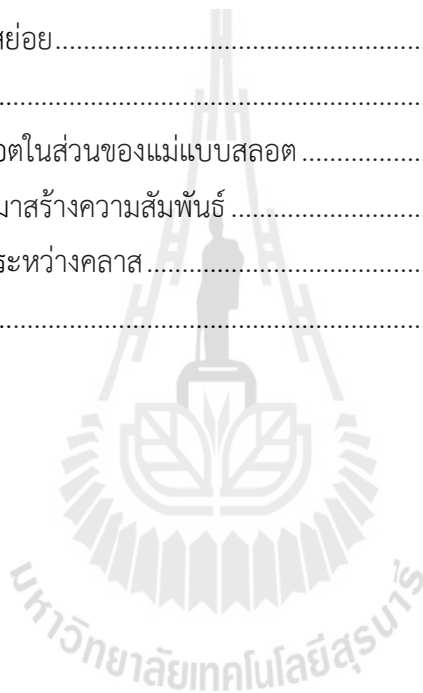
ภาพที่	หน้า
2.1 ตัวอย่างการวิเคราะห์โครงสร้างของวัสดุตามทฤษฎีไวยากรณ์ไม่พึงบริบท	13
2.2 ตัวอย่างการวิเคราะห์โครงสร้างของประโยคตามทฤษฎีไวยากรณ์ไม่พึงบริบท	13
2.3 กระบวนการพัฒนาออนโทโลยี (ที่มา Noy and McGuinness, 2003)	17
2.4 โครงสร้างของออนโทโลยีด้านการท่องเที่ยว (ในส่วนของที่พักและที่พัก).....	20
2.5 สมการความหมายของความพึงพอใจของลูกค้า	23
2.6 อิทธิพลของแรงจูงใจต่อความพึงพอใจในการท่องเที่ยว	24
2.7 เว็บไซต์ด้านการท่องเที่ยว TripAdvisor.com.....	27
3.1 วิธีการดำเนินงานวิจัย	30
3.2 โครงสร้างคลาสของออนโทโลยีด้านการท่องเที่ยวของเสียร์เพส พรานท์เนอร์ และบาชเลchner (Siorpaes, Prantner and Bachlechner, 2004).....	32
3.3 โครงสร้างคลาสของออนโทโลยีด้านการท่องเที่ยวที่สร้างขึ้นในงานวิจัยนี้	33
3.4 คะแนนที่แสดงระดับความพึงพอใจที่แตกต่างกัน	35
3.5 กฎไวยากรณ์ไม่พึงบริบทของระบบ	36
3.6 โครงสร้างของระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับ ธุรกิจท่องเที่ยวไทย.....	37
3.7 โครงสร้างของระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับ ธุรกิจท่องเที่ยวไทย.....	38
3.8 ตัวอย่างการทำงานของตัววิเคราะห์คำศัพท์	40
3.9 ตัวอย่างการทำงานของตัววิเคราะห์เชิงโครงสร้าง.....	40
3.10 ตัวอย่างการคำนวณระดับคะแนนคุณลักษณะของที่พัก	44
3.11 ตัวอย่างการคำนวณระดับคะแนนคุณลักษณะของที่พักแบบเป็นลำดับขั้น	44
3.12 การแบ่งกลุ่มสมาชิกตัวแปรระดับคะแนน	45
3.13 ตัวอย่างการอนุมานค่าผลลัพธ์ เมื่อค่าน้ำหนักของกฎเป็น 1 และ 0.5.....	49
3.14 ตัวอย่างการคำนวณระดับคะแนนในภาพรวมของเรื่องห้องพัก เมื่อคะแนนเรื่องเครื่องปรับอากาศ เป็น 4 คะแนน และเรื่องผ้าปูเตียงเป็น 2 คะแนน	50
3.15 ตัวอย่างการคำนวณระดับคะแนนในภาพรวมของทั้งโรงแรม เมื่อคะแนนเรื่องห้องพักเป็น 3.14 คะแนน เรื่องการบริการเป็น 2 คะแนน และเรื่องทำเลที่ตั้งเป็น 5 คะแนน	50
3.16 ตัวอย่างการแสดงผลความรู้ที่สกัดได้	52

สารบัญญภาพ (ต่อ)

ภาพที่	หน้า
3.17 หน้าจอหลักของโปรแกรม Protégé	55
4.1 หน้าจอหลักของระบบ.....	58
4.2 ส่วนการค้นหาข้อมูลที่ปัก.....	59
4.3 ส่วนการเข้าสู่ระบบและเมนูการใช้งานตามสิทธิของผู้ใช้.....	59
4.4 ส่วนเมนูลัดในการเข้าใช้งานระบบ	60
4.5 ส่วนแนะนำที่ปัก	61
4.6 ส่วนแสดงการส่งเสริมการขายของที่ปัก.....	61
4.7 ส่วนแสดงรูปภาพของที่ปัก.....	62
4.8 ส่วนแสดงการค้นหาที่ปักและเขียนบทวิจารณ์ วิธีที่ 1	63
4.9 ส่วนแสดงการค้นหาที่ปักและเขียนบทวิจารณ์ วิธีที่ 2	63
4.10 แผนภาพต้นไม้การจัดระดับความพึงพอใจของที่ปัก	65
4.11 ส่วนแสดงผลการสกัดความรู้จากคำวิจารณ์ ในกรณีมีคำวิจารณ์หลายเรื่อง	65
4.12 ผลการสกัดความรู้จากคำวิจารณ์ ในกรณีที่ระบบไม่สามารถสกัดได้.....	66
4.13 หน้าจอส่วนแสดงสารสนเทศทางภูมิศาสตร์ออนไลน์ของบริการสาธารณสุขรอบสถานที่ปัก	66
4.14 หน้าจอส่วนแสดงสารสนเทศทางภูมิศาสตร์ออนไลน์ของบริการอาหารและ เครื่องดื่มรอบสถานที่ปัก	67
4.15 หน้าจอส่วนแสดงสารสนเทศทางภูมิศาสตร์ออนไลน์ของบริการร้านค้ารอบสถานที่ปัก	68
4.16 หน้าจอส่วนแสดงสารสนเทศทางภูมิศาสตร์ออนไลน์ของบริการรถเข็นรอบสถานที่ปัก	68
4.17 หน้าจอส่วนแสดงสารสนเทศทางภูมิศาสตร์ออนไลน์ เมื่อคลิกที่จุดปักหนึ่งและปรับรัศมีของ ระยะทางการแสดงผลสิ่งอำนวยความสะดวกรอบสถานที่ปักเป็น 5 กิโลเมตร.....	69
4.18 แผนภูมิสรุปผลการประเมินความถูกต้องของการสกัดคุณลักษณะที่สำคัญของสถานที่ปัก.....	76
4.19 ผลการวิเคราะห์ค่าระดับคะแนนความสัมพันธ์ระหว่างคำวิเศษณ์และคำคุณศัพท์ (r)	77
ก.1 คลังคำศัพท์ของระบบ.....	120
ก.2 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 1	143
ก.3 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 2	144
ก.4 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 3	145
ก.5 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 4	146
ก.6 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 5	147

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
ข.1 หน้าจอหลักของโปรแกรม Protégé.....	149
ข.2 กล้องข้อความโต้ตอบกับผู้ใช้เพื่อสร้างหรือเปิดโครงการ.....	149
ข.3 โปรแกรม Protégé แสดงรายละเอียดของแท็บคลาส.....	150
ข.4 กล้องข้อความโต้ตอบกับผู้ใช้เพื่อบันทึกโครงการ.....	151
ข.5 หน้าจอแสดงการสร้างคลาส.....	151
ข.6 หน้าจอแสดงการสร้างคลาสย่อย.....	152
ข.7 หน้าต่างสล็อตฟอร์ม.....	153
ข.8 หน้าจอแสดงรายละเอียดสล็อตในส่วน of แม่แบบสล็อต.....	153
ข.9 หน้าต่างเลือกคลาสที่จะนำมาสร้างความสัมพันธ์.....	154
ข.10 ผลการสร้างความสัมพันธ์ระหว่างคลาส.....	154
ข.11 ผลการสร้างอินสแตนซ์.....	155



บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหาการวิจัย

ในปัจจุบัน การพัฒนาทางด้านเทคโนโลยีสารสนเทศและการสื่อสารมีความเจริญก้าวหน้าเป็นอย่างมาก ซึ่งเป็นผลทำให้เกิดกิจกรรมทางเศรษฐกิจและสังคมตลอดจนการเปลี่ยนแปลงอย่างกว้างขวาง ทั้งภายในประเทศและระดับนานาชาติ จากการพัฒนาของเทคโนโลยีสารสนเทศและการสื่อสารดังกล่าว ส่งผลให้อินเทอร์เน็ตได้เข้ามามีบทบาทในชีวิตประจำวันของผู้คนมากยิ่งขึ้น โดยทั่วโลกมีผู้ใช้งานอินเทอร์เน็ตสูงถึงประมาณ 1,800 ล้านคน (คิดเป็นร้อยละ 26.6 ของประชากรทั้งโลก) ในขณะที่ประเทศไทยก็มีผู้ใช้งานอินเทอร์เน็ตถึงประมาณ 16.1 ล้านคน (Miniwatts, www, 2009: 1) จากจำนวนการใช้งานอินเทอร์เน็ตที่สูงมากดังกล่าว ได้ส่งผลให้การดำเนินกิจกรรมทางธุรกิจแบบดั้งเดิมเปลี่ยนรูปแบบไปเป็นธุรกิจอิเล็กทรอนิกส์ (E-business) มากขึ้น

ธุรกิจอิเล็กทรอนิกส์ เป็นการทำธุรกรรมผ่านสื่ออิเล็กทรอนิกส์ในทุกช่องทาง เช่น การซื้อขายสินค้าและบริการ การโฆษณาผ่านสื่ออิเล็กทรอนิกส์ ไม่ว่าจะเป็นโทรศัพท์ โทรทัศน์ วิทยุ หรือแม้แต่อินเทอร์เน็ต โดยมีวัตถุประสงค์เพื่อลดค่าใช้จ่าย และเพิ่มประสิทธิภาพขององค์กร โดยการลดบทบาทองค์ประกอบทางธุรกิจ อาทิ อาคารประกอบการ คลังสินค้า รถขนส่งสินค้า และพนักงาน ทำให้ช่วยลดข้อจำกัดด้านสถานที่ ระยะทาง เวลา และแรงงานลงได้ (กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร, www, 2550: 1) ในปัจจุบันการประกอบธุรกิจอิเล็กทรอนิกส์และพาณิชย์อิเล็กทรอนิกส์ได้แพร่หลายมากขึ้นโดยดำเนินการบนเครือข่ายอินเทอร์เน็ต เช่น การยื่นแบบแสดงรายการเสียภาษีผ่านเว็บไซต์ของกรมสรรพากร การซื้อขายสินค้า การชำระเงินด้วยบัตรเครดิตผ่านอินเทอร์เน็ต การประมูลสินค้า การจองบัตรเข้าชมภาพยนตร์ การจองที่พักที่พักร้านอาหารและบริการอื่น ๆ ของแหล่งท่องเที่ยว หรือแม้แต่การติดตามการจัดส่งไปรษณีย์ลงทะเบียนผ่านเว็บไซต์ของบริษัท ไปรษณีย์ไทย จำกัด

ในการดำเนินการทางธุรกิจนั้น การสร้างความน่าเชื่อถือและสร้างความพึงพอใจให้แก่ลูกค้า (Customer Satisfaction) เป็นเรื่องที่ต้องพิจารณาอย่างยิ่ง เนื่องจากสถานการณ์การแข่งขันทางธุรกิจได้ทวีความรุนแรงมากขึ้น ในขณะที่จำนวนลูกค้ายังคงเท่าเดิม ผู้เป็นลูกค้าจึงมีทางเลือกมากขึ้นในการเลือกซื้อสินค้าหรือบริการ ด้วยเหตุนี้ผู้ประกอบการจึงพยายามสรรหาวิธีที่จะสร้างความพึงพอใจแก่ลูกค้าให้ได้มากที่สุด ทั้งต่อตัวสินค้า การบริการ และต่อองค์กร โดยเฉพาะในธุรกิจอิเล็กทรอนิกส์ที่ดำเนินการบนช่องทางอินเทอร์เน็ต เนื่องจากบางองค์กรในธุรกิจนี้ไม่มีสถานประกอบการในโลกความจริง การสร้างความพึงพอใจให้แก่ลูกค้าจึงเป็นเรื่องที่สำคัญมาก (กองประมวลผล และกองระบบสารสนเทศ, www, 2544: 1)

ในที่นี้ขอยกตัวอย่างธุรกิจด้านการท่องเที่ยว ปัจจุบันการดำเนินการขององค์กรต่าง ๆ ที่มีส่วนเกี่ยวข้องกับธุรกิจนี้ ต่างได้มีการปรับเปลี่ยนองค์กรของตนเองไปสู่ธุรกิจอิเล็กทรอนิกส์มากยิ่งขึ้น เพื่อเป็นแรงจูงใจให้ลูกค้าหรือนักท่องเที่ยวมาซื้อสินค้าหรือบริการของตน และในด้านของนักท่องเที่ยวนั้น ธุรกิจอิเล็กทรอนิกส์นี้มีส่วนช่วยอย่างมากในการอำนวยความสะดวกในด้านการบริการจองที่พักที่พักรถ หรือจองตั๋วเครื่องบิน และปัจจุบันยังมีส่วนสำคัญในการช่วยให้นักท่องเที่ยวสามารถสืบค้นข้อมูลด้านการท่องเที่ยว เพื่อนำมาใช้ประกอบการตัดสินใจเลือกแหล่งท่องเที่ยวให้ได้ตรงกับความต้องการอีกด้วย โดยวิธีที่นิยมใช้กันโดยทั่วไปคือ การตัดสินใจโดยใช้ข้อมูลจากบทวิจารณ์ของนักท่องเที่ยวที่เคยไปมาแล้ว ซึ่งแนวโน้มดังกล่าวได้สอดคล้องกับผลการศึกษาขององค์การส่งเสริมการท่องเที่ยวโลก (World Tourism Organization: WTO) ที่พบว่าปัจจัยสำคัญที่ส่งผลให้การท่องเที่ยวโลกเพิ่มสูงขึ้น คือ ระบบข้อมูลข่าวสารทันสมัยและก้าวหน้ามากขึ้น เพราะเมื่อนักท่องเที่ยวมีโอกาสรับทราบข้อมูลข่าวสารด้านการท่องเที่ยวมากขึ้น จะทำให้นักท่องเที่ยวมีความสนใจและมั่นใจในการเดินทางเพิ่มมากขึ้นตามไปด้วย (World Tourism Organization, 2007: 8-9)

อย่างไรก็ตาม ถึงแม้ว่าการจัดเก็บข้อมูลต่าง ๆ จากลูกค้าจะมีการปฏิบัติกันอยู่แล้วอย่างกว้างขวางในโลกธุรกิจอิเล็กทรอนิกส์ แต่การจะนำข้อมูลเหล่านั้นมาใช้ประโยชน์กลับเป็นเรื่องที่ยุ่งยากและต้องใช้เวลาานาน เนื่องจากข้อมูลที่จัดเก็บไว้นั้นเป็นข้อมูลดิบที่ยังไม่ได้ผ่านการคัดกรอง การประมวลผล หรือการสกัด เพื่อให้ได้ความรู้ที่สามารถนำไปใช้ประกอบการตัดสินใจได้ทันที ตัวอย่างเช่น เว็บไซต์สำหรับการท่องเที่ยวที่มีการจัดเก็บบทวิจารณ์ของนักท่องเที่ยวที่ได้ไปแหล่งท่องเที่ยวมาแล้ว เว็บไซต์ประเภทนี้ มีวัตถุประสงค์เพื่อให้ลูกค้าหรือนักท่องเที่ยวที่เคยไปท่องเที่ยวมาแล้ว ร่วมแสดงความคิดเห็นเกี่ยวกับแหล่งท่องเที่ยว ที่พัก หรือร้านอาหารในแหล่งท่องเที่ยว นั้น ๆ เพื่อให้ข้อมูลกับนักท่องเที่ยวคนอื่น ๆ ซึ่งยังมีความต้องการที่ไม่ชัดเจนว่าจะไปท่องเที่ยวที่ใด และต้องการค้นหาข้อมูลเพื่อนำไปใช้ในการประกอบการวางแผนท่องเที่ยวของตน เว็บไซต์ประเภทนี้จะเพียงให้ข้อมูลอันด้วยความนิยมของแหล่งท่องเที่ยวโดยรวม ที่นักท่องเที่ยวคนอื่นได้เคยแสดงความคิดเห็นด้วยวิธีการลงคะแนนเท่านั้น หากนักท่องเที่ยวคนใดต้องการรายละเอียดเพิ่มเติม จะต้องตามอ่านในทุก ๆ ความคิดเห็นเอง แล้วจึงวิเคราะห์หรือตัดสินใจว่าแหล่งท่องเที่ยว นั้นควรไปหรือไม่ เนื่องจากเว็บไซต์เป็นแบบตายตัว ข้อมูลที่ถูกจัดเก็บไว้จะเป็นข้อมูลดิบยังไม่ได้ผ่านการประมวลผล ทำให้เป็นข้อมูลที่ยังไม่ละเอียดเพียงพอสำหรับความต้องการนักท่องเที่ยว ดังนั้น งานวิจัยนี้จึงได้มุ่งพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทยขึ้นมา เพื่อใช้เป็นเครื่องมือที่ช่วยเพิ่มความสามารถในการสกัดเอาความรู้ที่เป็นประโยชน์ต่อการประกอบธุรกิจอิเล็กทรอนิกส์ และสนับสนุนให้ผู้ประกอบการสามารถเก็บและวิเคราะห์ข้อมูลเกี่ยวกับลูกค้าเพื่อนำไปใช้ประโยชน์ในด้านต่าง ๆ และสามารถแข่งขันทางการตลาดได้ อีกทั้งลูกค้ายังสามารถนำความรู้ในเรื่องที่สกัดได้ไปใช้ประกอบการตัดสินใจได้อย่างสะดวกและรวดเร็วอีกด้วย

ในปัจจุบัน ขั้นตอนวิธีการสกัดความรู้แบบอัตโนมัติที่มีอยู่จะทำงานอยู่บนพื้นฐานของคลังคำ (Corpus) หรือระดับของคำ เช่น อรรถาภิธาน (Thesaurus) และเวิร์ดเน็ต (Word Net) ทำให้สามารถสกัดคำได้เท่าที่มีอยู่ในคลังคำเท่านั้น หรือหากสกัดได้ก็มักเป็นคำที่ความหมายไม่ตรงกับที่ต้องการ (Choochart Haruechaiyasak, Prapass Srichaivattana, Sarawoot Kongyoung and Chaianun Damrongrat, 2004: 1) งานวิจัยนี้จึงได้นำออนโทโลยี (Ontology) มาประยุกต์ใช้ในการจัดเก็บคำศัพท์เพื่อสกัดและค้นคืนความรู้ เนื่องจากมีโครงสร้างการจัดเก็บข้อมูลคำต่าง ๆ อย่างเป็นลำดับขั้น ทำให้ทราบได้ว่าแต่ละคำมีความหมายและความสัมพันธ์กับคำอื่น ๆ อย่างไร ซึ่งหากนำมาประยุกต์ใช้ในการสกัดความรู้ จะทำให้ระบบสามารถประมวลความหมายหรือความสัมพันธ์ของคำได้อย่างตรงประเด็น ผลลัพธ์หรือความรู้ที่ได้จึงมีประสิทธิภาพมากกว่าการใช้คลังคำศัพท์หรือระดับของคำเป็นพื้นฐาน โดยงานวิจัยนี้ได้พัฒนาออนโทโลยีขึ้นมา เพื่อนำมาใช้กับการสกัดและค้นคืนความรู้แบบอัตโนมัติ โดยใช้ข้อมูลด้านการท่องเที่ยว อันได้แก่ แหล่งที่พักสำหรับนักท่องเที่ยว เป็นกรณีศึกษา

สำหรับระบบสกัดและค้นคืนความรู้เพื่อธุรกิจท่องเที่ยวไทยที่พัฒนาขึ้นมาในโครงการนี้ เป็นการผสมผสานระหว่างระบบเว็บเชิงความหมายและระบบสารสนเทศทางภูมิศาสตร์สำหรับการท่องเที่ยว ซึ่งระบบดังกล่าวจะช่วยนักท่องเที่ยวในการวางแผนการท่องเที่ยวโดยเฉพาะในเรื่องของที่พักให้เหมาะสมกับความต้องการของนักท่องเที่ยวมากที่สุด เนื่องจากได้รับคำแนะนำจากนักท่องเที่ยวที่เคยไปมาแล้วผ่านระบบสกัดความรู้ และได้รับสารสนเทศทางภูมิศาสตร์ซึ่งแสดงสิ่งอำนวยความสะดวกต่าง ๆ รอบที่พักนั้น ๆ นอกจากนี้นักท่องเที่ยวยังสามารถแสดงความคิดเห็นเกี่ยวกับที่พักต่าง ๆ เพื่อเป็นประโยชน์ต่อนักท่องเที่ยวกลุ่มอื่นได้อีกด้วย

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อเก็บรวบรวม และพัฒนาฐานความรู้ด้านการท่องเที่ยว เพื่อนำมาใช้ในการสกัดความรู้แบบอัตโนมัติ

1.2.2 เพื่อออกแบบและพัฒนาระบบสกัดความรู้แบบอัตโนมัติสำหรับธุรกิจท่องเที่ยวไทย

1.2.3 เพื่อออกแบบและพัฒนาระบบค้นคืนความรู้ผ่านสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย

1.3 สมมติฐานการวิจัย

1.3.1 ระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย สามารถประมวลความหมายจากบทวิจารณ์เกี่ยวกับสถานที่ที่นักท่องเที่ยวไปพักได้อย่างถูกต้อง โดยมีค่าความระลึกละมากกว่า 80%

1.3.2 ระบบสกัดและค้นคืนความรู้ที่พัฒนาขึ้นมา สามารถคำนวณระดับคะแนนความพึงพอใจต่อสถานที่พักของนักท่องเที่ยวได้อย่างถูกต้อง โดยมีความแตกต่างจากที่นักท่องเที่ยวกำหนดไม่เกิน 1 คะแนน

1.4 ข้อตกลงเบื้องต้น

1.4.1 งานวิจัยนี้ได้เลือกข้อมูลสถานที่พักสำหรับนักท่องเที่ยว มาเป็นข้อมูลตัวอย่างในการพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ เนื่องจากอุตสาหกรรมการท่องเที่ยวเป็นธุรกิจที่ทำรายได้สำคัญของประเทศไทย

1.4.2 ข้อมูลตั้งต้นในการพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย คือ บทวิจารณ์ของนักท่องเที่ยวที่เคยได้เข้ามาใช้บริการที่พักร้าน ๆ มาแล้ว

1.4.3 ผู้ใช้ต้องเข้าสู่ระบบเพื่อเขียนบทวิจารณ์เกี่ยวกับที่พักเป็นภาษาอังกฤษ

1.4.4 บทวิจารณ์เกี่ยวกับที่พักสำหรับนักท่องเที่ยวที่ป้อนเข้าสู่ระบบ จะต้องมีความถูกต้องตามหลักไวยากรณ์ภาษาอังกฤษ เช่น ใช้คำที่มีความหมาย และใช้เครื่องหมายมหัพภาค (.) ทุกครั้งเมื่อจบประโยคคำวิจารณ์ เป็นต้น

1.5 ขอบเขตของการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงประยุกต์ ซึ่งได้ออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย โดยมุ่งเน้นการพัฒนากระบวนการในการวิเคราะห์บทวิจารณ์อย่างละเอียดแบบอัตโนมัติ โดยระบบจะใช้ข้อมูลตั้งต้นเป็นประสบการณ์ของนักท่องเที่ยวที่ได้ไปสถานที่นั้น ๆ มาแล้ว และได้เข้าสู่ระบบเพื่อแสดงความคิดเห็นเกี่ยวกับสถานที่นั้น ๆ ข้อมูลดังกล่าวจะถูกนำไปใช้ในการปรับปรุงออนไลน์การท่องเที่ยวให้เหมาะสม และนำไปใช้ในการสกัดความรู้แบบอัตโนมัติต่อไป ทั้งนี้ความรู้ที่สกัดได้ทั้งหมดจะถูกจัดเก็บลงในฐานความรู้ของระบบแบบเรียลไทม์ นอกจากนี้ ระบบยังได้พัฒนาระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ ซึ่งแสดงสิ่งอำนวยความสะดวกต่าง ๆ รอบสถานที่นั้น ๆ เพื่อช่วยนักท่องเที่ยวในการวางแผนการท่องเที่ยว โดยเฉพาะในเรื่องของที่พักให้เหมาะสมกับความต้องการของนักท่องเที่ยวมากที่สุด

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1.6.1 ได้กระบวนการในการสกัดและค้นคืนความรู้แบบอัตโนมัติจากบทวิจารณ์ออนไลน์ของนักท่องเที่ยว

1.6.2 ได้ระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย

1.6.3 ผู้ประกอบธุรกิจอิเล็กทรอนิกส์

- สามารถนำระบบสกัดและค้นคืนความรู้แบบอัตโนมัติที่พัฒนาขึ้นไปใช้เป็นต้นแบบในการพัฒนาระบบอื่น ๆ เพื่อเก็บและวิเคราะห์ข้อมูลเกี่ยวกับลูกค้าที่เป็นประโยชน์ต่อธุรกิจ
- สามารถนำข้อมูลเหล่านั้นมาวิเคราะห์และพัฒนาสินค้าหรือบริการให้ตอบสนองความต้องการของลูกค้า เพื่อเปลี่ยนจากผู้บริโภคไปสู่การเป็นลูกค้าตลอดไป
- สามารถนำข้อมูลที่ได้จากการวิเคราะห์และประมวลผลแล้ว มาใช้ในการวางแผนเชิงนโยบายด้านการบริหารและการจัดการ

1.6.4 นักท่องเที่ยว

- สามารถนำความรู้ในเรื่องที่สกัดได้ไปใช้ประกอบการตัดสินใจได้อย่างสะดวกและรวดเร็ว
- สามารถนำสารสนเทศทางภูมิศาสตร์ออนไลน์ ซึ่งแสดงสิ่งอำนวยความสะดวกต่าง ๆ รอบสถานที่นั้น ๆ ไปช่วยในการวางแผนการท่องเที่ยวได้อย่างมีประสิทธิภาพ

1.7 คำอธิบายศัพท์

1.7.1 การสกัดและค้นคืนความรู้ (Knowledge Extraction and Retrieval)

หมายถึง กระบวนการในการค้นหาความรู้ที่เฉพาะเจาะจงจากคลังข้อความภาษาธรรมชาติ (Natural Language) หรือคลังข้อความที่ไม่มีโครงสร้าง โดยการวิเคราะห์ข้อมูลเพื่อจำแนกประเภทและความสัมพันธ์ของข้อมูล เพื่อให้ได้ความรู้ในเรื่องนั้น ๆ และนำความรู้ที่สกัดได้ไปใช้

1.7.2 ออนโทโลยี (Ontology)

หมายถึง โครงสร้างฐานความรู้ทางด้านใดด้านหนึ่งในขอบเขตที่สนใจ โดยใช้วิธีการจัดกลุ่มเพื่อแยกความแตกต่าง แล้วนำมาเข้าพวกเพื่อให้ง่ายต่อการสืบหาหรืออ้างอิง โดยมีลักษณะโครงสร้างแบบลำดับชั้น (Tree) และมีความสัมพันธ์กันระหว่างข้อมูล

1.7.3 ผู้เชี่ยวชาญ (Expert)

หมายถึง บุคคลที่มีความรู้หรือมีความเชี่ยวชาญทางด้านภาษาอังกฤษ

1.7.4 บทวิจารณ์ (Review)

หมายถึง ประสบการณ์หรือความคิดเห็นของนักท่องเที่ยวเกี่ยวกับสถานที่พักที่เคยไปมาแล้ว ซึ่งได้วิจารณ์ผ่านเว็บไซต์

1.7.5 คำวิจารณ์ (Comment)

หมายถึง ข้อความหรือประโยคที่วิจารณ์ถึงเรื่องใดเรื่องหนึ่งในบทความ

1.7.6 การท่องเที่ยว (Tourism)

หมายถึง การเดินทางไปยังแหล่งหรือสถานที่เป้าหมาย เพื่อวัตถุประสงค์อย่างใดอย่างหนึ่งหรือหลายอย่าง อาทิ เพื่อผ่อนคลายอิริยาบถจากงานประจำและสร้างความรื่นรมย์ของจิตใจ เพื่อศึกษาค้นหาคำตอบ หรือแลกเปลี่ยน ความรู้ และประสบการณ์ เพื่อจรรโลงใจและการโน้มน้าวใจให้เกิดความตระหนักและสำนึก เป็นต้น

1.7.7 การท่องเที่ยวออนไลน์ (E-tourism)

หมายถึง ธุรกิจด้านการท่องเที่ยวที่มีการดำเนินกิจกรรมทางธุรกิจผ่านสื่ออิเล็กทรอนิกส์ เช่น การประชาสัมพันธ์แหล่งท่องเที่ยว การจองที่พักที่พักรและบริการอื่น ๆ ผ่านเครือข่ายอินเทอร์เน็ต

1.7.8 ทฤษฎีไวยากรณ์ไม่พึ่งบริบท (Context Free Grammar)

หมายถึง กลุ่มของกฎไวยากรณ์ที่เขียนขึ้นมาโดยแสดงเป็นสัญลักษณ์ทางภาษา ซึ่งสามารถนำมาจัดกลุ่มและเรียงลำดับเข้าด้วยกันได้ และนำไปใช้ในการแปลภาษาแบบอัตโนมัติ



บทที่ 2

ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

ในบทนี้กล่าวถึง แนวคิดและงานวิจัยที่เกี่ยวข้องกับการวิจัยเรื่อง การพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย ซึ่งประกอบด้วย การสกัดและค้นคืนความรู้แบบอัตโนมัติ ทฤษฎีตัวแปลภาษา ทฤษฎีออนโทโลยี ระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ แนวคิดเกี่ยวกับความพึงพอใจของนักท่องเที่ยว และแนวคิดเกี่ยวกับบทวิจารณ์ออนไลน์ของนักท่องเที่ยว โดยมีรายละเอียดดังนี้

2.1 การสกัดและค้นคืนความรู้แบบอัตโนมัติ

งานวิจัยนี้ เป็นการออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย ดังนั้นในการออกแบบและพัฒนาระบบดังกล่าวจึงต้องศึกษาขั้นตอนวิธีและประโยชน์ของการสกัดและค้นคืนความรู้แบบอัตโนมัติที่มีอยู่ในปัจจุบัน เพื่อนำมาออกแบบและพัฒนาระบบให้มีความสมบูรณ์เหมาะสำหรับการนำไปใช้เพื่อการทดลองและศึกษาข้อมูลต่อไป

2.1.1 ความเป็นมาของการสกัดและค้นคืนความรู้แบบอัตโนมัติ

ข้อมูลข่าวสารต่าง ๆ ในปัจจุบันมีการเพิ่มขึ้นอย่างรวดเร็ว โดยส่วนใหญ่จะอยู่ในรูปแบบของเอกสารอิเล็กทรอนิกส์ ซึ่งเป็นการยากที่ผู้อ่านจะสามารถบริโภคข้อมูลข่าวสารต่าง ๆ เหล่านี้ได้ครบถ้วนภายในเวลาที่จำกัด ทำให้เกิดการสูญเสียโอกาสในการได้รับข้อมูลข่าวสารที่สำคัญได้ ดังนั้นกระบวนการของการสกัดความรู้แบบอัตโนมัติจึงเข้ามามีบทบาทในการนำเอกสารทั่วไป ซึ่งอยู่ในรูปแบบที่ไม่มีโครงสร้างหรือกึ่งโครงสร้าง มาประมวลผลให้อยู่ในรูปแบบของข้อมูลที่มีโครงสร้างชัดเจน

ประโยชน์ของการสกัดและค้นคืนความรู้แบบอัตโนมัติ ไม่เพียงแต่จะทำให้ผู้ใช้สามารถเข้าถึงและใช้งานความรู้นั้นได้สะดวกขึ้น แต่ยังสามารถนำความรู้ไปใช้ประมวลผลหรือตัดสินใจต่อได้ จึงทำให้ในปัจจุบันมีการนำความรู้เหล่านี้มาประยุกต์ใช้ในการวิเคราะห์และสร้างผลลัพธ์ที่เป็นประโยชน์มากขึ้นตามความต้องการของผู้ใช้ อาทิ การคำนวณข้อมูลทางสถิติ และการวิเคราะห์และแนะนำแนวทางแก้ปัญหาให้กับผู้ใช้ (อัศนีย์ ก่อตระกูล, 2550: 1-21)

สำหรับปัจจัยที่ส่งผลกระทบต่อประสิทธิภาพการทำงานของระบบสกัดและค้นคืนความรู้แบบอัตโนมัตินั้นมีหลายประการ ซึ่งปัจจัยที่สำคัญ ได้แก่ การประมวลเอกสารที่ไม่มีโครงสร้าง หรือเรียกว่า ภาษาธรรมชาติ การกำหนดรูปแบบที่ใช้สกัดหรือกรองข้อมูล และการบำรุงรักษาฐานความรู้ที่เกี่ยวข้อง (อัศนีย์ ก่อตระกูล, 2550: 1-1)

2.1.2 ขั้นตอนวิธีการสกัดและค้นคืนความรู้แบบอัตโนมัติ

ในปัจจุบันมีงานวิจัยเกี่ยวกับการสกัดและค้นคืนความรู้แบบอัตโนมัติจำนวนมาก ซึ่งแต่ละงานมีขั้นตอนวิธีที่แตกต่างกัน ทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ของแต่ละงานวิจัย เช่น งานวิจัยที่ทำงานอยู่บนพื้นฐานของคลังคำ ซึ่งมีการนำเทคนิคต่าง ๆ มาประยุกต์ใช้เพื่อให้ได้ผลลัพธ์ที่มีความถูกต้องมากขึ้น เช่น การนำโปรแกรมที่ใช้โครงสร้างการตัดสินใจแบบต้นไม้ (Decision Tree) มาช่วยในการสกัดคำภาษาไทยจากคลังคำศัพท์ และมีการระบุคำศัพท์ต่าง ๆ โดยใช้สูตรคำนวณ (Virach Somlertlamvanich, Tanapong Potipiti and Thatsanee Charoenporn, 2000) นอกจากนี้ยังมีการสร้างขั้นตอนวิธีที่เรียกว่าเอซีเคอี (Automatic Categorized Keyword Extraction: ACKE) เพื่อใช้ในการสกัดและค้นคืนคำสำคัญในภาษาไทยจากคลังข้อความที่มีการจัดแบ่งเป็นหมวดหมู่ โดยได้ทำการทดลองสกัดคำสำคัญจากบทความในหนังสือพิมพ์ (Choochart Haruechaiyasak et al., 2004) และโรมีโอ โอลิวาส เจเนโร และเพียทีนี (Romero, Olivas, Genero and Piattini, 2008) ได้ทำการพัฒนาเทคนิคที่ใช้ในการสกัดคำจากวารสารการทดลองด้านวิศวกรรมซอฟต์แวร์ (Empirical Software Engineering) โดยมีเป้าหมายเพื่อระบุคำตามกลุ่มคำศัพท์ที่สร้างจากความรู้ของผู้เชี่ยวชาญ (Wikiese) กลุ่มคำศัพท์นี้จะถูกปรับปรุงอยู่เสมอจากสารสนเทศที่สกัดได้และมีความเกี่ยวข้องกับหัวข้อวิจัยนั้น ซึ่งผลลัพธ์ที่สกัดได้จะถูกนำไปพัฒนาเป็นออนโทโลยีเพื่อใช้ในการสืบค้นเชิงความหมายต่อไป

นอกจากนี้ ยังมีงานวิจัยที่เกี่ยวข้องกับการสกัดและค้นคืนข้อมูลจากเอกสารในรูปแบบของเว็บไซต์ ซึ่งงานวิจัยส่วนใหญ่จะทำโดยการสร้างตัวแปลงเอกสารอัตโนมัติ (Wrapper) เพื่อแปลงเอกสารเอชทีเอ็มแอล (HTML) ให้อยู่ในรูปแบบที่ผู้ใช้กำหนดไว้ โดยการเรียนรู้จากตัวอย่างเอกสารที่ได้มีการกำกับไว้เป็นตัวอย่าง (Worapoj Peerawit, Warat Yingsaree and Asanee Kawtrakul, 2004) แต่วิธีการเช่นนี้ยังมีข้อจำกัดในเรื่องขอบเขตของข้อมูล เนื่องจากข้อมูลในขอบเขตที่แตกต่างกันในบางครั้งจะมีโครงสร้างเว็บที่มีรูปแบบแตกต่างกันด้วย จากการสำรวจพบว่า มีนักวิจัยพยายามที่จะแปลงโครงสร้างของเอกสารเอชทีเอ็มแอลให้อยู่ในรูปแบบต้นไม้ (Hutchatai Chanlekha and Asanee Kawtrakul, 2001) และยังมีนักวิจัยที่มุ่งเน้นไปยังการสกัดตารางออกจากเอกสารเอชทีเอ็มแอล (Tem Smitinand, 2001) เพื่อวิเคราะห์ตารางและสกัดเอาข้อมูลออกมา (Tem Smitinand, 2001; Chalermpon Sirigayon, Hutchatai Chanlekha and Asanee Kawtrakul, 2004; Aurawan Imsombut, Mukda Suktarachan, Warat Yingsaree and Asanee Kawtrakul, 2005; Jirawan Chareonsuk, Thana Sukvakree and Asanee Kawtrakul, 2005)

สุภารัตน์ วุฒิเลิศเจริญวงศ์ (Suparat Wuttildercharoenwong, 2009) ได้นำเสนอระบบในการสกัดและค้นคืนข้อมูลผู้เชี่ยวชาญด้านการเกษตร โดยแบ่งออกเป็นการสกัดข้อมูลผู้เชี่ยวชาญจากหลากหลายเว็บไซต์ที่มีความน่าเชื่อถือ และการรวบรวมข้อมูลเพื่อการค้นหาผู้เชี่ยวชาญ

ในเวลาอันสั้น งานวิจัยนี้ได้ประยุกต์ใช้เทคโนโลยีการสกัดข้อมูลและภาษาสำหรับการสร้างกฎ (Simple Rule Language: SRL) ในการสร้างกฎเพื่อการสกัดข้อมูล ซึ่งระบบที่พัฒนาขึ้นนี้ ยังจำกัดการสกัดและค้นคืนข้อมูลจากเว็บไซต์ที่มีโครงสร้างแบบกึ่งโครงสร้าง (Semi-structure) สำหรับการประเมินระบบจาก 5 เว็บไซต์ ความถูกต้องของการสกัดข้อมูลคิดเป็นร้อยละ 91.5

จากการทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้องกับการสกัดและค้นคืนความรู้แบบอัตโนมัติพบว่า การสกัดและค้นคืนความรู้จำเป็นอย่างยิ่งกับเอกสารประเภทไม่มีโครงสร้างหรือกึ่งโครงสร้าง เช่น ข้อมูลบนเว็บไซต์ ทำให้มีนักวิจัยจำนวนมากที่มุ่งเน้นเพื่อจะพัฒนาขั้นตอนวิธีที่มีประสิทธิภาพในการที่จะนำความรู้ออกมาจากเอกสารเหล่านั้น เพื่อช่วยให้ผู้ใช้สามารถเข้าถึงและใช้งานความรู้นั้นได้สะดวกขึ้น อีกทั้งยังนำความรู้ไปใช้ประมวลผลหรือตัดสินใจต่อได้โดยง่าย ซึ่งในงานวิจัยนี้เป็นการนำเอาบทวิจารณ์ออนไลน์ของนักท่องเที่ยวที่เป็นข้อมูลแบบไม่มีโครงสร้าง มาสกัดเป็นความรู้เพื่อใช้สนับสนุนการตัดสินใจของนักท่องเที่ยวได้อย่างสะดวกและรวดเร็วยิ่งขึ้น โดยมุ่งเน้นไปที่การพัฒนาขั้นตอนวิธีโดยใช้ทฤษฎีตัวแปลภาษาเข้ามาช่วย ซึ่งจะกล่าวถึงในหัวข้อถัดไป

2.2 ทฤษฎีตัวแปลภาษา

งานวิจัยนี้ เกี่ยวข้องกับการพัฒนากระบวนการในการวิเคราะห์ความหมายจากบทวิจารณ์ของนักท่องเที่ยวคนอื่น ๆ ที่ได้ไปแหล่งท่องเที่ยวแล้วมาสกัดเป็นความรู้ และนำความรู้ที่สกัดได้ทั้งหมดมาจัดเก็บลงในฐานความรู้ของระบบ ดังนั้น งานวิจัยนี้จึงเลือกศึกษาทฤษฎีตัวแปลภาษา เพื่อนำใช้ในการสกัดความรู้แบบอัตโนมัติ นั้น โดยมีรายละเอียดดังต่อไปนี้

2.2.1 ที่มาและกระบวนการแปลภาษา

การประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) เป็นศาสตร์สาขาย่อยสาขาหนึ่งของปัญญาประดิษฐ์ เพื่อให้คอมพิวเตอร์สามารถรับรู้และเข้าใจภาษามนุษย์ได้ เช่น การรับข้อมูลเข้าจากการพิมพ์ผ่านแป้นพิมพ์ หรือเสียงพูดผ่านไมโครโฟน หรือจากการอ่านเข้าระบบผ่านแฟ้มข้อมูล แล้วจัดการตัดทอนเอาเฉพาะที่มีความหมายสมบูรณ์เข้าเก็บในคอมพิวเตอร์ สิ่งที่มีความหมายจากการตัดทอนนี้ อาจเป็นคำสั่งที่จะส่งไปเรียกใช้ฐานข้อมูลโครงสร้างไวยากรณ์ทางภาษา เพื่อใช้ในการแปลภาษา หรือชุดของรหัสแอสกี (ASCII) เทียบได้กับเสียงพูดที่ผ่านการแปลงเป็นตัวอักษร ลักษณะงานของการประมวลผลภาษาธรรมชาติต่าง ๆ นี้ ส่วนใหญ่จะมีกระบวนการที่คล้ายคลึงกัน แต่จะมีความแตกต่างที่รูปแบบของการแสดงผล (มหาวิทยาลัยเกษตรศาสตร์, www, 2544: 1)

ในการสกัดความรู้แบบอัตโนมัติ จำเป็นต้องมีการประมวลผลภาษาธรรมชาติ โดยใช้ทฤษฎีทางด้านตัวแปลภาษา (Parser) เข้ามาช่วยในการตรวจสอบความถูกต้องของรูปประโยคในภาษาธรรมชาติหรือข้อความแบบไร้โครงสร้าง เนื่องจากในข้อความภาษาธรรมชาตินี้ คำที่เขียนหรือ

เปล่งออกมาเมื่อนำมารวมกันจะเกิดความหมายขึ้น คอมพิวเตอร์จึงจำเป็นต้องมีกระบวนการต่าง ๆ เพื่อจะได้เข้าใจความหมายของแต่ละประโยค ซึ่งมีกระบวนการที่สำคัญ (บุญเจริญ ศิริเนาวกุล, 2551: 386-387) ดังนี้

- การประมวลผลหน่วยคำ (Morphological Processing) เป็นการหารูปแบบของคำที่แท้จริงในประโยค ซึ่งอาจมีการแต่งเติมเพื่อให้เกิดความหมายเพิ่มเติมจากคำเดิม เช่น การเติม “ed” ให้กับคำกริยา การเติม “ing” การเติม “tion” หรือการเปลี่ยนรูปกริยา เป็นต้น การประมวลผลหน่วยคำจะเน้นหนักไปที่การตัดคำ (Word Segmentation) หรือการพิจารณาขอบเขตของคำในประโยค โดยในภาษาอังกฤษใช้ช่องว่างเป็นเครื่องหมายบอกขอบเขตของคำ

- การประมวลผลทางไวยากรณ์ (Syntactic Processing) เป็นการประมวลผลเพื่อหารูปแบบ ตรวจสอบ และสร้างไวยากรณ์ของภาษาที่ทำการประมวลผล อาทิ ประโยค “The room is clean.” ในแง่ของไวยากรณ์ภาษาอังกฤษ โครงสร้างของไวยากรณ์จะเป็น คำนำหน้านาม (DET: The) + นาม (NOUN: room) + กริยาช่วย (AUX: is) + คุณศัพท์ (ADJ: clean) โดยทั่วไปแล้วการประมวลผลไวยากรณ์จะใช้โครงสร้างต้นไม้นี้เป็นตัวบอกโครงสร้างไวยากรณ์ของประโยค

- การประมวลผลทางความหมาย (Semantic Processing) เป็นการหา และการวิเคราะห์ความหมายของประโยค ซึ่งหาได้มาจากความหมายของคำ และความสัมพันธ์ของคำในทางความหมาย หากสามารถกำหนดโครงสร้างทางความหมายที่เป็นไปได้ทั้งหมดออกมา จะสามารถนำมาวิเคราะห์ว่าประโยคใดมีความหมายผิดหรือไม่เป็นจริงได้

- การประมวลผลบทสนทนา (Discourse Processing) เป็นการหาความสัมพันธ์ของความหมายของคำระหว่างประโยค ในการวิเคราะห์ความหมายของคำบางคำในประโยค บางครั้งต้องพิจารณาความหมายของประโยคอื่นประกอบด้วย ตัวอย่างเช่น คำว่า “It” ของประโยค “It is very good.” หากไม่พิจารณาจากประโยคอื่นประกอบ ก็จะไม่ทราบว่า “It” ในที่นี้หมายถึงอะไร

- การประมวลผลความนัย (Pragmatic Processing) ในการหาความหมายที่แท้จริงของบางประโยค อาจต้องพิจารณาบริบทของข้อความทั้งข้อความประกอบด้วย เพราะในบางครั้งความหมายที่แท้จริงอาจจะไม่ได้มีความหมายเช่นเดียวกับคำที่แสดงไว้ในประโยคก็ได้ เช่น “Is it time to break up ?” อาจจะมี ความหมายในเชิงของการถามวันเวลา หรืออาจเป็นการบอกกล่าวให้พักผ่อนก็ได้ ขึ้นอยู่กับบริบทของข้อความในบทความนั้น

จากกระบวนการทั้งห้าที่กล่าวมาข้างต้น หากสามารถทำได้ทุกกระบวนการจะทำให้การประมวลผลภาษาธรรมชาติมีความถูกต้องอย่างยิ่ง แต่อย่างไรก็ตาม ในการวิจัยและพัฒนาที่พบโดยส่วนใหญ่จะใช้เพียง 3 กระบวนการแรกเท่านั้นคือ การประมวลผลหน่วยคำ การประมวลผลทางไวยากรณ์ และการประมวลผลทางความหมาย เนื่องจากกระบวนการถัดมาคือ การประมวลผลบทสนทนา และการประมวลผลความนัย ทำได้ยาก ใช้เวลานาน และไม่ได้ได้รับความถูกต้องเพิ่มขึ้นมาอย่างมีนัยสำคัญ

มากเท่าที่ควร ดังนั้นงานวิจัยส่วนใหญ่จึงมุ่งไปที่การพัฒนากระบวนการประมวลผลหน่วยคำ การประมวลผลทางไวยากรณ์ และการประมวลผลทางความหมาย ให้มีประสิทธิภาพมากที่สุดเป็นหลัก ซึ่งในแต่ละกระบวนการจะกล่าวอย่างละเอียดในหัวข้อถัดไป

2.2.2 การประมวลผลหน่วยคำ (Morphological Processing)

การประมวลผลหน่วยคำ (Morphological Processing) เป็นการศึกษาารูปแบบและโครงสร้างเกี่ยวกับการสร้างคำจากหน่วยที่เล็กที่สุดทางไวยากรณ์ของภาษา ซึ่งก็คือหน่วยคำ (Morpheme) (Jurafsky and Martin, 1998: 57) การประมวลผลหน่วยคำมีบทบาทสำคัญคือทำให้ทราบหลักการกฎเกณฑ์เกี่ยวกับโครงสร้างของประโยค เพราะเมื่อศึกษาโครงสร้างของการสร้างคำแล้ว จะต้องศึกษาหน้าที่ของคำ (Function of Word) ทั้งในระดับประโยค (Sentence Level) ระดับวลี (Phrase Level) และระดับคำ (Word Level)

การศึกษาระดับหน่วยคำ เป็นการศึกษาการสร้างคำใหม่จากรากคำเดิม (Derivational Morphology) ซึ่งอาจมีการแต่งเติมเพื่อให้เกิดความหมายเพิ่มเติมจากคำเดิม ทั้งนี้ในการศึกษาระดับหน่วยคำก็เพื่อหารูปแบบของรากคำเดิมในประโยค โดยในการสร้างคำใหม่จากรากคำเดิมมีหลายวิธีดังต่อไปนี้

- การเปลี่ยนรูปคำนาม หรือคำกริยา เช่น การเปลี่ยนรูปคำนามเอกพจน์ “mouse” เป็นคำนามพหูพจน์ “mice” หรือการเปลี่ยนรูปคำกริยาช่อง 1 “eat” เป็นกริยาช่อง 2 “ate” หรือเป็นกริยาช่อง 3 “eaten” เป็นต้น

- การเติมคำนำหน้า (Prefix) เช่น การเติม “un” “im” หรือ “in” ให้กับคำคุณศัพท์ (Adjective) หรือคำกริยาวิเศษณ์ (Adverb) การเติม “Re” “Dis” หรือ “Mis” ให้กับคำกริยา และการเติม “Pre” “Bi” หรือ “Tri” ให้กับคำนาม ตัวอย่างเช่น การเติม “un” ให้กับคำว่า “unhappy” “unsuitable” หรือ “uncountable” เป็นต้น

- การเติมคำต่อท้าย (Suffix) เช่น การเติม “ed” ให้กับคำกริยา การเติม “ing” หรือ “tion” ให้กับคำนาม และการเติม “s” ให้กับคำพหูพจน์ ซึ่งตัวอย่างของการเติมคำต่อท้าย อาทิ คำเติมหรือรากศัพท์คือ คำว่า “walk” สามารถเติมคำต่อท้ายได้เป็น “walked” “walking” หรือ “walks”

อย่างไรก็ตาม งานทางด้านภาษาที่พบในปัจจุบัน การสร้างคำใหม่อาจเกิดจากการเขียนในรูปแบบที่แตกต่างกันเองของมนุษย์ ซึ่งอาจเกิดจากความง่ายหรือความสะดวก เพื่อละคำให้สั้นลง อาทิ คำว่า “Air Conditioning” อาจเขียนเป็น “a/c” “air con” “air conditioning” “air conditioned” หรือ “air conditions” เป็นต้น

ในการประมวลผลหน่วยคำจะเริ่มต้นจากการตัดคำ (Word Segmentation) หรือการพิจารณาขอบเขตของคำในประโยค โดยในภาษาอังกฤษใช้ช่องว่างเป็นเครื่องหมายบอกขอบเขตของคำ

และในการหาคำเดิมหรือรากศัพท์ จะมีกระบวนการหรือเทคนิคที่หลากหลาย แต่ในงานวิจัยนี้ สำหรับการหารากศัพท์ของการเติมคำนำหน้าหรือคำต่อท้าย จะทำโดยการตัดคำนำหน้าหรือคำต่อท้ายที่พบทั้ง ส่วนการรากศัพท์จากคำที่เปลี่ยนรูป จะอาศัยคลังคำศัพท์ทั่วไป (Dictionary) เข้ามาช่วย ในขณะที่การหารากศัพท์จากการใช้คำย่อ จะอาศัยคลังคำเหมือนที่เก็บไว้ในฐานความรู้เป็นหลัก ซึ่งคลังคำศัพท์ทั่วไป และคลังเหมือนจะกล่าวถึงอย่างละเอียดในบทที่ 3

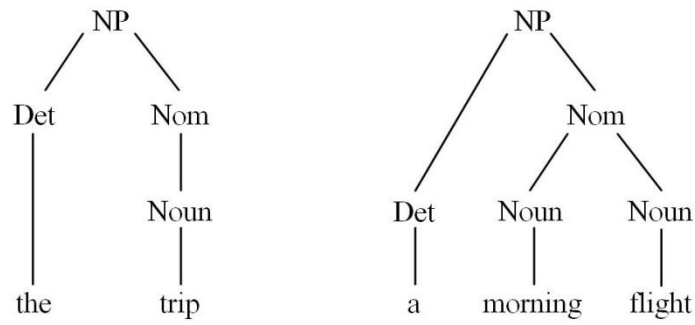
2.2.3 การประมวลผลทางไวยากรณ์ (Syntactic Processing)

ในปัจจุบัน เทคนิคที่ใช้ในการประมวลผลทางไวยากรณ์มีอยู่เป็นจำนวนมาก แต่เทคนิคที่นิยมนำมาใช้มากที่สุดคือ ทฤษฎีไวยากรณ์ไม่พึ่งบริบท เทคนิคนี้เป็นการสร้างหลักไวยากรณ์ขึ้นมาใช้ในการอธิบายโครงสร้างของภาษา และความสัมพันธ์ระหว่างโครงสร้างต่าง ๆ ที่ประกอบขึ้นในประโยคนั้น ๆ โดยไม่พิจารณาในเชิงความหมายของประโยค (ยีน ภูววรรณ และชัยยงค์ วงศ์ชัยสุวัฒน์, 2535: 58) และปรับโครงสร้างของภาษาที่เดิมมีลักษณะเชิงเส้นให้เป็นโครงสร้างต้นไม้ แล้ววิเคราะห์หา ลักษณะของคำที่เหมาะสมในประโยค โดยการกำหนดกลุ่มของกฎไวยากรณ์ที่แสดงเป็นสัญลักษณ์ทางภาษา จากนั้นจึงนำไปใช้ในการแปลภาษาแบบอัตโนมัติ ทั้งนี้กลุ่มของกฎไวยากรณ์ที่สร้างขึ้นจะประกอบด้วย 2 ส่วน คือ (1) นอนเทอร์มินอล (Non-terminal) หมายถึง ชนิดหรือหมวดคำศัพท์ และ (2) เทอมินอล (Terminal) หมายถึง คู่ของนอนเทอร์มินอล หรือคำศัพท์เดี่ยวในวลีหรือประโยคที่ต้องการวิเคราะห์ โดยมีกระบวนการวิเคราะห์ดังตัวอย่าง (Fry, 2004: 1-4)

ตัวอย่างที่ 1 กำหนดให้วลีที่ต้องการวิเคราะห์คือ “the trip” และ “a morning flight” และมีการกำหนดกลุ่มของกฎไวยากรณ์เพื่อใช้ในการวิเคราะห์วลีไว้ดังนี้

1. NP -> Det Nom
2. Nom -> Noun | Noun-Nom
3. Det -> a | the
4. Noun -> flight | trip | morning

จากกฎข้างต้น นอนเทอร์มินอลจะปรากฏทางด้านซ้ายมือ ได้แก่ NP Nom Det และ Noun ส่วนเทอร์มินอลจะปรากฏทางด้านขวามือ ได้แก่ คู่ของนอนเทอร์มินอล เช่น “Det Nom” และคำศัพท์เดี่ยวในวลีที่ต้องการวิเคราะห์ เช่น คำว่า “a” หรือ “the” สำหรับกระบวนการวิเคราะห์จะพิจารณาถึงคำศัพท์แต่ละคำในวลีว่าตรงกับนอนเทอร์มินอลใดตามกฎที่สร้างขึ้น แล้วจึงนำนอนเทอร์มินอลนั้นมาเขียนหรือนิยามใหม่ในระดับถัดไปตามโครงสร้างต้นไม้ จากนั้นจะพิจารณาที่ระดับปัจจุบันว่าสามารถนิยามใหม่ตามกฎใดได้อีก เมื่อทำการวิเคราะห์โครงสร้างวลีทั้งสองตามกลุ่มของกฎไวยากรณ์ที่สร้างขึ้น จะได้ผลเป็นโครงสร้างต้นไม้ดังภาพที่ 2.1

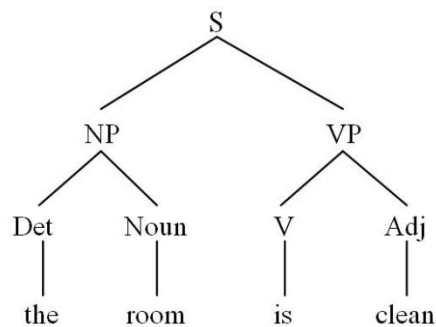


ภาพที่ 2.1 ตัวอย่างการวิเคราะห์โครงสร้างของวลีตามทฤษฎีไวยากรณ์ไม่พึ่งบริบท

ตัวอย่างที่ 2 กำหนดให้ประโยคที่ต้องการวิเคราะห์คือ “the room is clean” และมีกำหนดกลุ่มของกฎไวยากรณ์เพื่อใช้ในการวิเคราะห์ประโยคไว้ดังนี้

1. S → NP VP
2. NP → Det Noun
3. VP → V Adj
4. Det → the
5. Noun → room
6. V → is
7. Adj → clean

จากกฎข้างต้น นอนเทอมินอลปรากฏทางด้านซ้ายมือ ได้แก่ S NP VP Det Noun V และ Adj ส่วนเทอมินอลปรากฏทางด้านขวามือ ได้แก่ คู่ของนอนเทอมินอล “NP VP” หรือ “Det Noun” และคำศัพท์เดี่ยวในประโยคที่ต้องการวิเคราะห์ เช่น คำว่า “the” และ “room” สำหรับกระบวนการวิเคราะห์จะพิจารณาเช่นเดียวกับตัวอย่างที่ 1 และเมื่อทำการวิเคราะห์โครงสร้างประโยคตามกลุ่มของกฎไวยากรณ์ที่สร้างขึ้น จะได้ผลดังภาพที่ 2.2



ภาพที่ 2.2 ตัวอย่างการวิเคราะห์โครงสร้างของประโยคตามทฤษฎีไวยากรณ์ไม่พึ่งบริบท

2.2.4 การประมวลผลทางความหมาย (Semantic Processing)

การประมวลผลทางความหมายคือการวิเคราะห์หาความหมายของประโยค โดยประโยคที่กำลังพิจารณาอาจจะเขียนถูกต้องตามหลักไวยากรณ์ แต่มีความหมายกำกวมหรือเป็นความหมายที่เป็นไปไม่ได้ก็ได้ เช่น ประโยค “The room is clever” เป็นประโยคที่ถูกต้องตามไวยากรณ์ แต่ไม่มีความหมายในความเป็นจริง คำว่ามีความหมายในวิจัยนี้จะพิจารณาดังนี้ หากค่านามของประโยคจากบทวิจารณ์ที่กำลังพิจารณาอยู่ไม่เป็นคำคุณลักษณะที่สำคัญของที่พัก จะถือว่าประโยคดังกล่าวไม่มีความหมายในงานวิจัยนี้ แต่หากค่านามดังกล่าวเป็นคำคุณลักษณะที่สำคัญของที่พัก จะพิจารณาคำคุณศัพท์และคำวิเศษณ์ที่ขยายค่านามดังกล่าวว่ามีความหมายหรือไม่ หากมีจะตรวจสอบต่อไปว่ามีความหมายในระดับความพึงพอใจระดับใด ตัวอย่างเช่น บทวิจารณ์เขียนว่า “Location is great. You can actually walk to the Silom area in 10 minutes. Other places can be accessible by skytrain or taxi, very convenient. Bedsheet were clean but a bit old. Reception staffs were friendly and polite.” บทวิจารณ์ดังกล่าวปรากฏค่านามที่เป็นคำคุณลักษณะที่สำคัญของที่พัก สามคำ คือ Location, Bedsheet และ Reception ดังนั้นประโยคที่ถือว่ามีความหมายในงานนี้ คือ “Location is great.” “Bedsheet were clean but a bit old” และ “Reception staffs were friendly and polite.” ซึ่งประโยคที่มีความหมายทั้งสามประโยคดังกล่าว จะถูกตรวจสอบและให้คะแนนระดับความพึงพอใจ ซึ่งประโยค “Location is great.” และ “Reception staffs were friendly and polite.” จะถูกให้คะแนนระดับความพึงพอใจเป็นระดับดี ในขณะที่ประโยค “Bedsheet were clean but a bit old” จะถูกให้คะแนนระดับความพึงพอใจเป็นระดับปานกลาง สำหรับรายละเอียดของการให้คะแนนระดับความพึงพอใจจะกล่าวรายละเอียดในบทที่ 3

จากข้อมูลข้างต้นสามารถสรุปได้ว่า ตัวแปลภาษา เป็นกระบวนการที่สำคัญในการวิเคราะห์ความหมายในประโยคแบบไร้โครงสร้าง หรือบทวิจารณ์ของนักท่องเที่ยวได้ โดยมีกระบวนการที่สำคัญได้แก่ การประมวลผลหน่วยคำ การประมวลผลทางไวยากรณ์ และการประมวลผลทางความหมาย ซึ่งตัวแปลภาษานับเป็นเครื่องมือที่สำคัญในการพัฒนาระบบสกัดและค้นคืนความรู้แบบอัตโนมัติ

2.3 ทฤษฎีออนโทโลยี

งานวิจัยนี้ ได้นำออนโทโลยีมาประยุกต์ใช้ในการจัดเก็บคำศัพท์เพื่อสกัดและค้นคืนความรู้ โดยนำมาประยุกต์ใช้กับธุรกิจอิเล็กทรอนิกส์ในด้านการท่องเที่ยว อันได้แก่ แหล่งที่พัก เป็นกรณีศึกษาดังนั้น งานวิจัยนี้จึงเลือกศึกษาทฤษฎีออนโทโลยี เพื่อเป็นประโยชน์ต่อการนำไปพัฒนาระบบให้มีความสมบูรณ์มากที่สุด มีรายละเอียดดังนี้

2.3.1 นิยามของออนโทโลยี

ในสาขาที่เกี่ยวข้องกับปัญญาประดิษฐ์และวิศวกรรมความรู้ ออนโทโลยีเป็นคำที่ใช้กันอย่างแพร่หลาย โดยมีผู้ให้นิยามไว้ดังนี้

กริบเบอร์ (Gruber, 1993: 199) ได้นิยามว่า “ออนโทโลยี เป็นการให้รายละเอียดในเชิงมโนภาพ (*An ontology is an explicit specification of a conceptualization.*)” ซึ่งเป็นโครงสร้างของความรู้ในเชิงแนวความคิดซึ่งบรรยายความรู้อย่างมีขอบเขตโดยใช้คลาส (Classes) หรือมโนทัศน์ (Concepts) ความสัมพันธ์ (Relations) หน้าที่ (Functions) สัจพจน์ (Axioms) และสมาชิกของคลาส (Instances) ดังนั้นออนโทโลยีถือเป็นฐานความรู้ที่ถูกสร้างขึ้นโดยมีขอบเขต (Domain) ที่จำกัด และมีความสำคัญในแง่ของการเป็นศูนย์กลางของความรู้ ซึ่งสามารถถูกใช้งานร่วมกันได้ และจากการที่ออนโทโลยีสามารถใช้งานร่วมกันได้นั้น ทำให้มีข้อดีในแง่ของการนำกลับมาใช้ใหม่ และยังสามารถเพิ่มเติมหรือเปลี่ยนแปลงองค์ประกอบต่าง ๆ ได้ในภายหลัง

สวาทเอาท์ พาทิล ไนท์ และรัส (Swartout, Patil, Knight and Russ, 1996: 138) ได้นิยามว่า “ออนโทโลยีเป็นระบบคำศัพท์ที่มีความสัมพันธ์ในเชิงความหมาย และใช้เป็นโครงร่างพื้นฐานในการอธิบายความรู้เฉพาะด้าน (*An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base.*)”

สมชาย ปรากฏการเจริญ (2548) อธิบายว่า ออนโทโลยีเป็นศาสตร์ของการจัดหมวดหมู่ของสิ่งหนึ่ง ๆ ในขอบเขตงานที่สนใจ (Application Domain) อย่างมีระบบ มีกฎเกณฑ์เงื่อนไข แสดงความเป็นจริงเท็จได้อย่างถูกต้อง ภายใต้ความเห็นชอบยอมรับของบุคคลที่เกี่ยวข้อง (Expert Consensus) โดยหมวดหมู่ของข้อมูลเหล่านั้นต้องตอบคำถามทุก ๆ ข้อได้ นั่นคือความรู้ (Concept) หรือความหมายของสิ่งที่มีอยู่ (Existing) ได้ถูกกล่าวอ้างไว้ครบถ้วน หากมนุษย์และเครื่องจักรสามารถเข้าถึงความรู้ต่าง ๆ ที่ถูกจัดหมวดหมู่ครอบคลุมทุกคำถามที่ต้องการจากสิ่งที่ปรากฏได้อย่างถูกต้อง

ปัจจุบันออนโทโลยีเป็นเทคโนโลยีที่ถูกนำไปประยุกต์ใช้งานในหลายสาขา ซึ่งอาจมีวัตถุประสงค์หลักในการใช้งานแตกต่างกันออกไป เช่น การจัดการองค์ความรู้ (Knowledge Management) พาณิชยกรรมอิเล็กทรอนิกส์ (E-Commerce) หรือเว็บเซอร์วิส (Web Services) เป็นต้น ทั้งนี้ เพื่อใช้ในการสร้างความเข้าใจ หรือเป็นการสร้างข้อตกลงร่วมกันระหว่างผู้ใช้งานในขอบเขตเนื้อหาใดด้านหนึ่งที่สนใจ

ในการพัฒนาออนโทโลยีสำหรับธุรกิจท่องเที่ยวออนไลน์ ปัญหาที่พบในปัจจุบันคือ ไม่มีการกำหนดมาตรฐานคำจำกัดความที่เป็นมาตรฐานร่วมกันในการแสดงถึงลักษณะของผลิตภัณฑ์ และการบริการด้านการท่องเที่ยวที่ชัดเจน ซึ่งแสดงให้เห็นถึงการขาดความเชื่อมโยงกันของข้อมูลด้านการท่องเที่ยว ปัญหาที่กล่าวมานี้สามารถแก้ไขได้ด้วยการสร้างออนโทโลยี และการให้คำจำกัดความของคำต่าง ๆ (Waralak V. Siricharoen, 2007: 1275) จึงมีการพัฒนาเว็บเชิงความหมายโดยอาศัย

ออนโทโลยีในการกำหนดขอบเขตข้อมูล และทำให้ระบบมีความสามารถในการใช้ข้อมูลร่วมกันเพื่อพัฒนาธุรกิจท่องเที่ยวออนไลน์ (Roopa, Mladen and Nalin, 2007: 1) และมีการสำรวจการใช้งานเว็บไซต์ด้านการท่องเที่ยวที่เป็นเว็บเชิงความหมาย โดยอาศัยความสามารถของออนโทโลยี เพื่อศึกษาถึงการใช้ประโยชน์และปัญหาจากการใช้งานเว็บไซต์ที่เกิดขึ้น (Hepp, Siorpaes and Bachlechner, 2006: 1)

ในงานด้านการสกัดและค้นคืนความรู้ นั้น ฐานข้อมูลความรู้เกี่ยวกับคำศัพท์หรือออนโทโลยีเป็นทรัพยากรความรู้เบื้องหลังที่มีความสำคัญ และช่วยให้การทำงานของระบบมีประสิทธิภาพมากยิ่งขึ้น

2.3.2 ประเภทของออนโทโลยี

ประเภทของออนโทโลยีตามขอบเขตของการประยุกต์ใช้งานเป็นสามประเภทดังนี้ (สมชาย ปราการเจริญ, 2548)

1) ออนโทโลยีแบบสาธารณะ (Generic) ออนโทโลยีนี้เป็นจุดรวมของโครงสร้างของความรู้ในเชิงกว้าง มีความเป็นนามธรรมค่อนข้างสูง (Abstract) เป็นความรู้ในเชิงมหัพภาพ สามารถถูกเรียกใช้งานคุณลักษณะต่าง ๆ ของออนโทโลยีได้โดยไม่ต้องสร้างใหม่ (Reuse) ออนโทโลยีสาธารณะนี้มีชื่อเรียกได้หลายลักษณะ เช่น ออนโทโลยีระดับบน (Upper Ontology) ออนโทโลยีแกนหลัก (Core Ontology) ออนโทโลยีอ้างอิง (Reference Ontology) ตัวอย่างของออนโทโลยีประเภทนี้ได้แก่ ไซค์อัปเปอร์ (Cyc-Upper) ซึ่งอธิบายถึงสิ่งต่าง ๆ แยกเป็นประเภท ๆ แต่ไม่ได้บรรยายรายละเอียดย่อยไว้

2) ออนโทโลยีสำหรับระบบงานหนึ่ง ๆ (Domain Oriented Ontology) เป็นออนโทโลยีที่จัดทำขึ้นโดยมีรายละเอียดครอบคลุมเฉพาะในระบบงานหนึ่งๆ มีการรวบรวมคุณลักษณะทั่วไปที่สามารถใช้ร่วมกัน (Generalization) ระหว่างกิจกรรมต่าง ๆ (Task) ภายในระบบงานนั้น ๆ

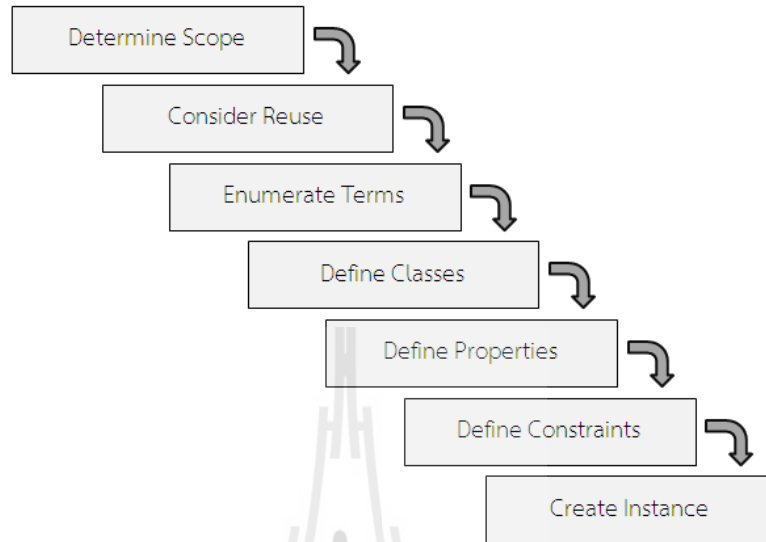
3) ออนโทโลยีสำหรับกิจกรรม (Task Oriented Ontology) เป็นออนโทโลยีที่พัฒนาขึ้นเพื่อตอบสนองการทำงานของกิจกรรมย่อย ๆ (Task) โดยอาศัยคุณลักษณะจากออนโทโลยีสำหรับระบบงานหนึ่ง ๆ มาช่วยบรรยายลักษณะเชิงสาธารณะ

2.3.3 กระบวนการพัฒนาออนโทโลยี

กระบวนการพัฒนาออนโทโลยี สามารถแบ่งออกได้เป็น 7 ขั้นตอนดังแสดงในภาพที่ 2.3 โดยกระบวนการดังกล่าวอธิบายได้ดังนี้ (Noy and McGuinness, 2003)

1) ระบุขอบเขตและวัตถุประสงค์ (Determine Scope) เป็นการกำหนดวัตถุประสงค์ของการพัฒนาออนโทโลยีซึ่งจะมีผลกระทบต่อารออกแบบ การประเมินผล และการนำกลับมาใช้ใหม่ของออนโทโลยี

2) พิจารณาการนำออนโทโลยีที่กลับมาใช้ใหม่ (Consider Reuse) เป็นการให้ความสำคัญกับนิยามต่าง ๆ ที่จำเป็นต้องใช้ เพื่อนำเข้าสู่ระบบโดยอาศัยเครื่องมือสำหรับช่วยในการพัฒนา



ภาพที่ 2.3 กระบวนการพัฒนาออนโทโลยี (ที่มา Noy and McGuinness, 2003)

3) พิจารณารายละเอียดของพจน์ (Enumerate Terms) เป็นขั้นตอนพิจารณารายละเอียดทั้งหมดว่าต้องการสื่อถึงเรื่องอะไรและมีคุณสมบัติอย่างไร โดยไม่ต้องคำนึงถึงความซ้ำซ้อนกันของแนวคิด และคุณสมบัติ

4) ขั้นตอนการกำหนดคลาส (Define Classes) คลาสหมายถึงแนวคิด (Concept) ที่อยู่ในเรื่องที่สนใจ (Domain) ประกอบด้วยส่วนประกอบต่าง ๆ โดยในการพัฒนาลำดับของคลาสมีวิธีการอยู่หลายวิธี แต่วิธีที่นิยมใช้ได้แก่

ก. การพัฒนาแบบบนลงล่าง (Top – Down) โดยเป็นการกำหนดแนวคิดทั่ว ๆ ไปของเรื่องที่สนใจและลำดับของแนวคิดก่อน แล้วค่อยแบ่งหมวดหมู่ของคลาส

ข. การพัฒนาแบบล่างขึ้นบน (Bottom – Up) เป็นการระบุคลาสส่วนใหญ่ก่อนแล้วจึงนำมาจัดกลุ่มให้เป็นแนวคิดใหญ่

ค. การพัฒนาแบบผสม (Combination) เป็นการผสมผสานระหว่างวิธีการพัฒนาแบบบนลงล่างและวิธีการแบบล่างขึ้นบน โดยจะทำเฉพาะแนวคิดที่สำคัญก่อน แล้วค่อยทำการจัดหมวดหมู่ของคลาส

5) การกำหนดคุณสมบัติของคลาส (Define Properties) โดยมีการกำหนดประเภทให้กับคุณสมบัติของคลาสและต้องพิจารณาว่าคลาสมีคุณสมบัติแบบง่าย (Simple Properties) หรือมีคุณสมบัติแบบซับซ้อน (Complex Properties)

6) กำหนดเงื่อนไขให้กับคุณสมบัติ (Define Constraints) เป็นการกำหนดเงื่อนไข (Facet) ให้กับคุณลักษณะ (Attribute หรือ Slots) ซึ่งแต่ละคุณลักษณะจะมีเงื่อนไขที่แตกต่างกัน โดยต้องพิจารณาจำนวนที่เป็นไปได้ของคุณลักษณะ (Slot Cardinality) เรื่องที่สนใจ (Domain) และช่วงของคุณลักษณะ (Slot Range)

7) สร้างสมาชิกของคลาส (Create Instance) หรือตัวข้อมูลในออนโทโลยี

2.3.4 ภาษาที่ใช้ในการพัฒนาออนโทโลยี

ภาษาที่ใช้ในการพัฒนาออนโทโลยีในเชิงกรอบความรู้ประกอบด้วย อาร์ดีเอฟ (RDF) อาร์ดีเอฟเอส (RDFS) และอวาล์ (OWL) ในงานวิจัยนี้ผู้วิจัยเลือกพัฒนาโดยใช้ภาษาอวาล์เนื่องจากภาษาอวาล์สามารถอธิบายความสัมพันธ์และคุณสมบัติของคลาสได้และสามารถถ่ายทอดคุณสมบัติ (Inherit) ให้แก่คลาสย่อยหรือคลาสลูกได้ดังนั้นภาษาอวาล์จึงสามารถอธิบายความจริงได้มากกว่าภาษาอาร์ดีเอฟหรืออาร์ดีเอฟเอส

อวาล์ (OWL: Web Ontology Language) สร้างขึ้นโดยเวิร์ลด์ไวด์เว็บคอนซอร์เทียม (W3C) เป็นภาษาออนโทโลยีสำหรับเว็บเชิงความหมาย (Semantic Web) แบ่งออกเป็น 3 ภาษาด้วยกันคือ อวาล์ ไลท์ (OWL Lite) อวาล์ ดี-แอล (OWL DL) และอวาล์ ฟูล (OWL Full) (McGuinness and Harmelen, 2004) ในงานวิจัยเป็นการพัฒนาออนโทโลยีอาหารไทยเพื่อที่จะชดเชยวัตถุดิบในการประกอบอาหารซึ่งต้องมีการใช้การอนุมานความรู้ซึ่งภาษาอวาล์ที่มีความสามารถในการให้เหตุผลได้คือ อวาล์ ดี-แอล ความรู้ที่ใช้ในการพัฒนาอวาล์ ดี-แอล (Horridge, 2011) มีดังนี้

ส่วนประกอบของอวาล์ประกอบด้วย สมาชิกของคลาส (Individuals) คุณสมบัติ (Properties) และคลาส (Classes) ดังนี้

2.3.4.1 สมาชิกของคลาส (Individuals หรือ Instances) คือ สิ่งที่น่าสนใจในเรื่องใดเรื่องหนึ่ง กล่าวอีกนัยหนึ่งคือ ตัวข้อมูลที่ใส่ลงไปเอนออนโทโลยี

2.3.4.2 คุณสมบัติ (Properties) คือความสัมพันธ์ที่เชื่อมระหว่างสมาชิกของคลาสหรือสิ่งสองสิ่ง

2.3.4.3 คลาส (Classes) คือเซตที่มีสมาชิก สามารถจัดลำดับเป็นซูเปอร์คลาสหรือคลาสแม่ (Superclass) กับคลาสย่อยหรือคลาสลูก (Subclass) ได้

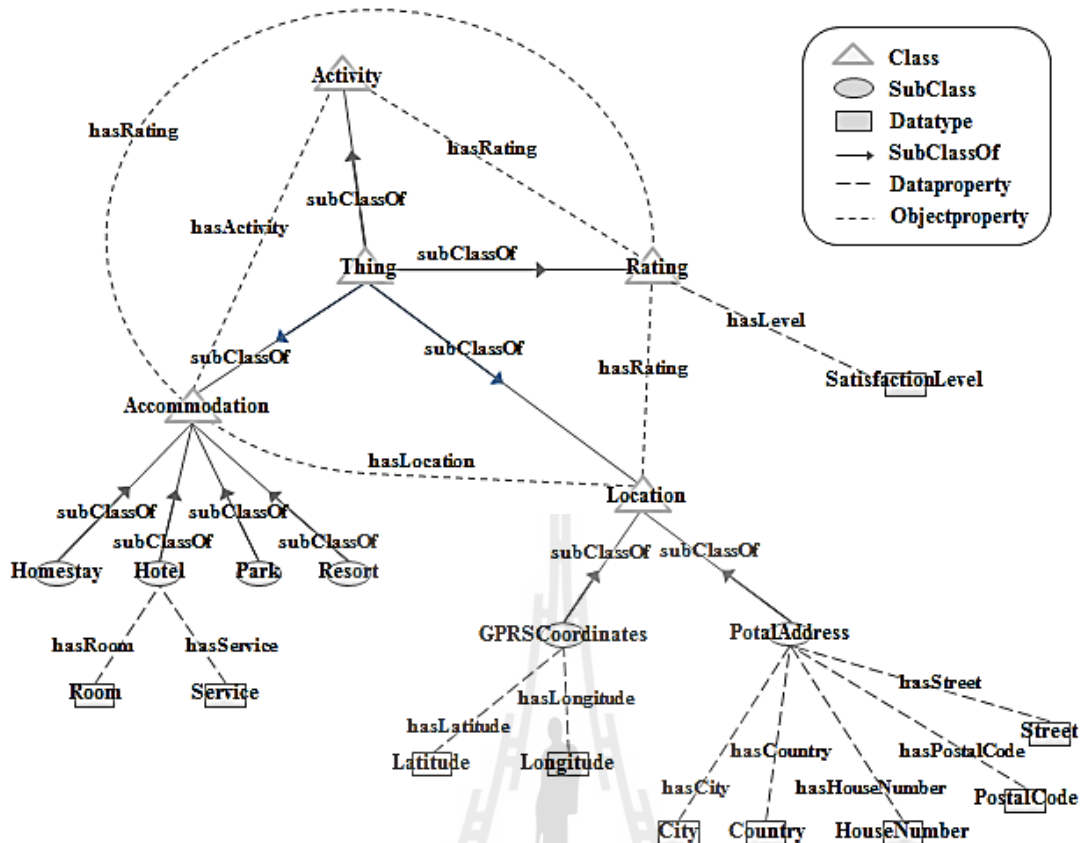
2.3.5 งานวิจัยที่นำออนโทโลยีเข้ามาช่วยในการสกัดและค้นคืนความรู้

งานวิจัยที่นำออนโทโลยีเข้ามาช่วยในการสกัดและค้นคืนความรู้ เช่น การศึกษาและพัฒนา ระบบการสกัดและค้นคืนความรู้จากเอกสารเว็บ โดยการสร้างออนโทโลยีเกี่ยวกับชีวประวัติของศิลปิน และนำมาใช้ในการระบุความรู้และอธิบายความสัมพันธ์ระหว่างคำจากฐานคำศัพท์ เพื่อนำมาสร้าง อัตชีวประวัติของศิลปินที่ผู้ใช้สนใจตามรูปแบบที่ได้กำหนดไว้ (Alani et al., 2003)

โปพอฟ และคณะ (Popov et al., 2003) ได้ทำการสกัดข้อมูลโดยใช้การจัดการความรู้ และสารสนเทศ (Knowledge and Information Management Platform: KIM) ซึ่งเป็นรูปแบบในการจัดทำดัชนีเชิงความหมาย (Semantic Indexing) การทำบรรณนิเทศน์ (Annotation) และการค้นคืนข้อมูล โดยงานวิจัยนี้ได้รวมเอาการสกัดข้อมูลที่ใช้รูปแบบการทำวิศวกรรมข้อความกับเว็บเชิงความหมาย ซึ่งเป็นตัวแทนของความรู้ และมีการนำออนโทโลยีมาประยุกต์ใช้ในการให้รายละเอียดเกี่ยวกับชนิดและชื่อของเอนทิตีโดยการตรวจสอบกับไวยากรณ์ และใช้ฐานความรู้ในการสร้างการค้นหบบรรณนิเทศน์ตลอดจนการจดจำความสัมพันธ์ระหว่างเอนทิตี เพื่อใช้ในการสกัดและค้นคืนความรู้ได้อย่างถูกต้องแม่นยำยิ่งขึ้น

บาร์โกลินี และคณะ (Bartolini et al., 2006) ได้นำเสนอวิธีการที่สนับสนุนการสกัดข้อมูลโดยใช้ออนโทโลยีสำหรับแหล่งข้อมูลมัลติมีเดีย โดยเฉพาะจากบัญชีรายชื่อผลิตภัณฑ์ (Product Catalogues) ซึ่งมีการนำออนโทโลยีหลักมาใช้ในกระบวนการสกัด และในขณะเดียวกันก็สร้างหรือขยายออนโทโลยีที่สามารถใช้งานได้ด้วย ผลที่ได้จากกระบวนการสกัด คือ ความรู้ที่สกัดได้จากข้อความรวมทั้งรูปภาพในบัญชีรายชื่อผลิตภัณฑ์ อาทิ ข้อมูลรายละเอียดผลิตภัณฑ์ ซึ่งความรู้นี้สามารถนำไปพัฒนาต่อยอดใช้กับบริการอื่น ๆ ได้

งานวิจัยนี้ได้พัฒนาออนโทโลยีสถานที่พักในด้านการท่องเที่ยว เพื่อนำมาใช้จัดเก็บคำศัพท์สำหรับการสกัดและค้นคืนความรู้ โดยใช้ข้อมูลด้านการท่องเที่ยวในส่วนของแหล่งที่พักเป็นกรณีศึกษา ดังแสดงโครงสร้างในภาพที่ 2.4 ในเบื้องต้นได้นำออนโทโลยีที่พัฒนาขึ้นมาใช้เพื่อสกัดความรู้จากบทวิจารณ์ของนักท่องเที่ยวในธุรกิจท่องเที่ยวออนไลน์ อย่างไรก็ตาม การนำออนโทโลยีและวิธีการสกัดและค้นคืนความรู้ที่มีอยู่มาประยุกต์ใช้ในงานด้านต่าง ๆ จำเป็นต้องมีการปรับเปลี่ยนกระบวนการให้เหมาะสมกับงานนั้น ๆ เพื่อให้การสกัดและค้นคืนความรู้แบบอัตโนมัติมีความถูกต้องและมีประสิทธิภาพ



ภาพที่ 2.4 โครงสร้างของออนโทโลยีด้านการท่องเที่ยว (ในส่วนที่เกี่ยวข้องกับสถานที่พัก)

2.4 ระบบสารสนเทศทางภูมิศาสตร์ออนไลน์

ระบบสกัดและค้นคืนความรู้เพื่อธุรกิจท่องเที่ยวไทยที่พัฒนาขึ้นมาในโครงการนี้ เป็นการผสมผสานระหว่างระบบเว็บเชิงความหมายและระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับการท่องเที่ยว ซึ่งระบบดังกล่าวจะช่วยนักท่องเที่ยวในการวางแผนการท่องเที่ยวโดยเฉพาะในเรื่องของที่พักให้เหมาะสมกับความต้องการของนักท่องเที่ยวมากที่สุด เนื่องจากได้รับคำแนะนำจากนักท่องเที่ยวที่เคยไปมาแล้วผ่านระบบสกัดความรู้ และได้รับสารสนเทศทางภูมิศาสตร์แบบออนไลน์ซึ่งแสดงสิ่งอำนวยความสะดวกต่าง ๆ รอบที่พักนั้น ๆ โดยความหมาย องค์ประกอบ และเทคโนโลยีที่นำมาใช้ในระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ มีรายละเอียดดังต่อไปนี้

2.4.1 ความหมายของระบบสารสนเทศทางภูมิศาสตร์ออนไลน์

ระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ (Online Geographic Information System: Online GIS) คือ ระบบสารสนเทศทางภูมิศาสตร์ (Chang, 2007) ซึ่งแสดงผลผ่านเว็บไซต์ โดยเป็นกระบวนการทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ด้วยระบบคอมพิวเตอร์และแสดงผลผ่านเว็บไซต์ด้วยการกำหนดข้อมูลและสารสนเทศที่มีความสัมพันธ์กับตำแหน่งในเชิงพื้นที่ เช่น ที่อยู่ สัมพันธ์กับ

ตำแหน่งพิกัดในแผนที่ (เส้นรุ้ง เส้นแวง) ซึ่งสามารถอ้างอิงถึงตำแหน่งที่มีอยู่จริงบนพื้นโลกได้ ข้อมูลและแผนที่ในระบบสารสนเทศทางภูมิศาสตร์ออนไลน์เป็นระบบสารสนเทศที่อยู่ในรูปของฐานข้อมูลที่มีส่วนสัมพันธ์กับข้อมูลเชิงพื้นที่ (Spatial Data) ซึ่งรูปแบบและความสัมพันธ์ของข้อมูลเชิงพื้นที่ทั้งหลายจะสามารถนำมาวิเคราะห์และแสดงผลด้วยระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ได้ เช่น ตำแหน่งของสิ่งอำนวยความสะดวกบริเวณแหล่งท่องเที่ยว จำนวนนักท่องเที่ยวในแต่ละบริเวณของแหล่งท่องเที่ยว ฯลฯ ข้อมูลเหล่านี้เมื่อปรากฏบนแผนที่ในเว็บไซต์ จะทำให้สามารถแปลและสื่อความหมายใช้งานได้ง่าย

2.4.2 องค์ประกอบของระบบสารสนเทศทางภูมิศาสตร์ออนไลน์

องค์ประกอบหลักของระบบสารสนเทศทางภูมิศาสตร์ แบ่งออกเป็น 5 ส่วนใหญ่ ๆ คือ อุปกรณ์คอมพิวเตอร์ (hardware) โปรแกรม (software) ข้อมูล (data) บุคลากร (people) และขั้นตอนการทำงาน (methods) โดยมีรายละเอียดของแต่ละองค์ประกอบดังต่อไปนี้

1. อุปกรณ์คอมพิวเตอร์ คือ เครื่องคอมพิวเตอร์รวมถึงอุปกรณ์ต่อพ่วง เช่น ดิจิทัล-เซอร์ สแกนเนอร์ เครื่องพิมพ์ เพื่อใช้ในการนำเข้าข้อมูล ประมวลผล และแสดงผล โดยที่คอมพิวเตอร์ของผู้ใช้งานระบบสารสนเทศทางภูมิศาสตร์ออนไลน์จะต้องสามารถเชื่อมต่อกับเครือข่ายอินเทอร์เน็ตได้เพราะจำเป็นจะต้องเข้าถึงและจัดการข้อมูลผ่านเว็บไซต์

2. โปรแกรม คือรหัสคำสั่งที่ทำให้คอมพิวเตอร์ทำงาน ประกอบด้วยฟังก์ชัน การทำงาน และเครื่องมือที่จำเป็นต่าง ๆ สำหรับนำเข้าและปรับแต่งข้อมูล จัดการฐานข้อมูล เรียกค้น วิเคราะห์ และ จำลองภาพ ซึ่งในงานวิจัยนี้ได้ใช้โปรแกรม ได้แก่ มายซีคิวล (MySQL) สำหรับจัดเก็บฐานข้อมูล กูเกิ้ลแมพเอพีไอ (Google Map APIs) สำหรับการเข้าถึงแผนที่แบบออนไลน์ ฯลฯ

3. ข้อมูล คือข้อมูล ที่จะใช้ในระบบสารสนเทศทางภูมิศาสตร์ออนไลน์และถูกจัดเก็บในรูปแบบของฐานข้อมูล ตัวอย่างของข้อมูล เช่น ข้อมูลชื่อ ประเภท ตำแหน่งพิกัด ของสิ่งอำนวยความสะดวก ต่าง ๆ ซึ่งข้อมูลเหล่านี้จะถูกนำไปประมวลผลสำหรับแสดงสิ่งอำนวยความสะดวกรอบสถานที่ที่กำหนดผ่านเว็บไซต์

4. บุคลากร คือ ผู้ปฏิบัติงานซึ่งเกี่ยวข้องกับระบบสารสนเทศภูมิศาสตร์ออนไลน์ อาทิ ผู้นำเข้าข้อมูล ช่างเทคนิค ผู้ดูแลระบบฐานข้อมูล ผู้เชี่ยวชาญสำหรับวิเคราะห์ข้อมูล

5. วิธีการหรือขั้นตอนการทำงาน คือ วิธีการที่นำเอาระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ไปใช้งาน เช่น ในงานวิจัยนี้ได้นำระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ไปใช้เพื่อแสดงตำแหน่งสิ่งอำนวยความสะดวกรอบสถานที่ท่องเที่ยว

2.4.3 ภูเก็ตแมพเอพีไอ

ภูเก็ตแมพเอพีไอ คือ ช่องทางในการเชื่อมต่อเพื่อเข้าใช้แผนที่ของบริษัทภูเก็ต ซึ่งเป็นเครื่องมือที่สำคัญมากที่สุดในในการพัฒนาระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ ภูเก็ตแมพเอพีไอจะมี ส่วนต่อประสานโปรแกรมประยุกต์หลายหมวด อาทิ หมวดการจัดการแผนที่ หมวดการควบคุมต่าง ๆ (เช่น ควบคุมการขยายและตำแหน่งของแผนที่) หมวดส่วนซ้อนทับบนแผนที่ (เช่น มาร์กเกอร์แสดง ตำแหน่งบนแผนที่ การวาดรูปทับซ้อนบนแผนที่) หมวดด้านบริการต่าง ๆ (เช่น การหาเส้นทางระหว่าง สถานที่ การหาระยะทาง) โดยปกติแล้วการเข้าถึงภูเก็ตแมพเอพีไอจะต้องเข้าผ่านภาษาจาวาสคริปต์ (Java Script) ซึ่งในการพัฒนาจริงจะต้องมีการเข้าใช้ฐานข้อมูลเพื่อเข้าถึงตำแหน่งพิกัดต่าง ๆ ด้วย ภาษาพีเอชพี ดังนั้นเทคโนโลยีเอแจ็กซ์ (AJAX) จึงเป็นสิ่งจำเป็นที่จะทำให้สามารถใช้งานฐานข้อมูลจาก จาวาสคริปต์ได้

2.5 แนวคิดเกี่ยวกับความพึงพอใจของนักท่องเที่ยว

งานวิจัยนี้ ได้นำบทวิจารณ์ของนักท่องเที่ยวมาใช้เป็นกรณีศึกษา เนื่องจากมีแนวคิดที่ว่า ความพึงพอใจของนักท่องเที่ยวเป็นสิ่งสำคัญ และมีอิทธิพลอย่างยิ่งต่อธุรกิจการท่องเที่ยว ดังนั้น งานวิจัยนี้จึงศึกษา แนวคิดเกี่ยวกับความพึงพอใจของลูกค้า และงานวิจัยที่เกี่ยวข้องกับความพึงพอใจของลูกค้า เพื่อศึกษา ปัจจัยที่มีอิทธิพลต่อการเลือกมาท่องเที่ยวของนักท่องเที่ยว และนำไปพัฒนาการแสดงผลของระบบให้ สอดคล้องกับความต้องการของผู้ใช้ระบบมากที่สุด มีรายละเอียดดังนี้

2.5.1 ความหมายและความสำคัญของความพึงพอใจของนักท่องเที่ยว

ในการดำเนินธุรกิจไม่ว่าประเภทใด การสร้างความน่าเชื่อถือและความพึงพอใจให้กับลูกค้าเป็นเรื่องที่ควรพิจารณาอย่างยิ่ง เนื่องจากความพึงพอใจของลูกค้าถือเป็นปัจจัยสำคัญในการแข่งขันเพื่อช่วงชิงความได้เปรียบทางธุรกิจ จากสภาวะเศรษฐกิจโลกซบเซาที่ส่งผลกระทบต่อผลประกอบการของธุรกิจ อีกทั้งสถานการณ์การแข่งขันที่ทวีความรุนแรงมากขึ้น ผู้ประกอบการธุรกิจจึงพยายามสรรหาวิธีที่จะสร้างความพึงพอใจแก่ลูกค้าให้ได้มากที่สุดทั้งต่อองค์กรและการบริการ ปัจจุบันผู้ประกอบการจึงมุ่งเน้นการดำเนินธุรกิจเพื่อตอบสนองความต้องการของลูกค้าเป็นหลัก (กอง ประมวลผล และกองระบบสารสนเทศ, www, 2544: 1)

โดยศิริวรรณ เสรีรัตน์ และคณะ (2538: 11) ได้กล่าวถึงความพึงพอใจของลูกค้าว่า “เป็นระดับความรู้สึกของลูกค้าที่มีผลจากการเปรียบเทียบระหว่างผลประโยชน์จากคุณสมบัติผลิตภัณฑ์หรือการทำงานของผลิตภัณฑ์กับการคาดหวังของบุคคล”

ณัฐพัชร์ ล้อประดิษฐ์พงษ์ (2549: 26) สรุปความหมายของความพึงพอใจของลูกค้าว่า “คือ ความรู้สึกของลูกค้าทั้งที่เกิดความพึงพอใจและและไม่พึงพอใจ อันเกิดจากการเปรียบเทียบ

ระหว่างสิ่งที่ได้รับจากการใช้สินค้าหรือบริการ (Perceived Performance) กับความคาดหวัง (Expectation)” ดังภาพที่ 2.5

$$\text{Customer Satisfaction} = \text{Expectation} - \text{Perceived Performance}$$

ภาพที่ 2.5 สมการความหมายของความพึงพอใจของลูกค้า

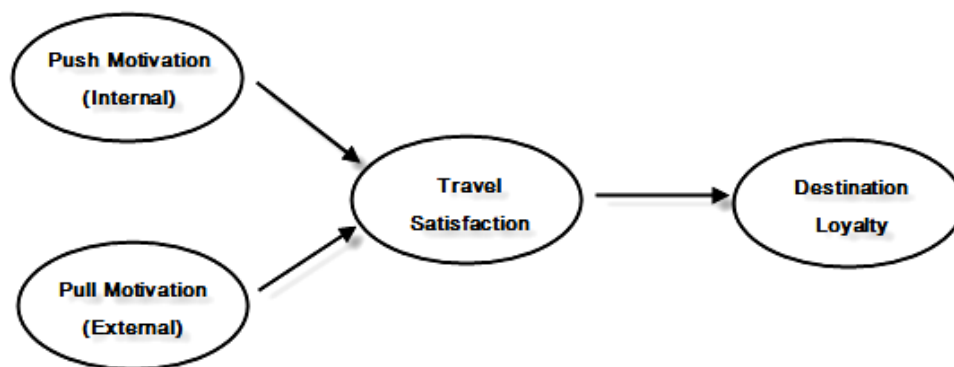
นอกจากนี้ ฌ็องพัซร์ ล็องประดิษฐ์พงษ์ ยังอธิบายเพิ่มเติมว่า ความภักดีของลูกค้า คือทัศนคติของลูกค้าที่มีต่อสินค้าและบริการที่นำไปสู่ความสัมพันธ์ในระยะยาว เป็นการเหนี่ยวรั้งลูกค้าไว้กับองค์กร พฤติกรรมที่ภักดีของลูกค้าเป็นผลมาจากการที่ลูกค้ามีความพึงพอใจมาก่อน จากนั้นจึงสะสมจนกลายมาเป็นความภักดี ดังนั้น การที่จะทำให้ลูกค้ามีความภักดีได้ต้องเริ่มต้นจากการสร้างความพึงพอใจให้เกิดขึ้นอย่างต่อเนื่อง องค์กรไม่สามารถสร้างความภักดีของลูกค้าให้เกิดขึ้นได้หากลูกค้าไม่พึงพอใจ

ความภักดีและความพึงพอใจของลูกค้ามีความแตกต่างกันตรงที่ความพึงพอใจมีลักษณะเป็นพลวัต (Dynamic) กล่าวคือ ความพึงพอใจของลูกค้ามักเปลี่ยนแปลงเสมอ อาจเพิ่มขึ้นหรือลดลงได้ตลอดเวลา ในขณะที่ความภักดีของลูกค้าจะมีความเสถียร (Static) มากกว่า เนื่องจากความภักดีของลูกค้าเป็นความรู้สึกประทับใจในตัวสินค้า ตราสินค้า ตลอดจนความสัมพันธ์ที่ตระหว่งธุรกิจกับลูกค้าหรือพนักงานกับลูกค้า ซึ่งมักเปลี่ยนแปลงช้ากว่าความพึงพอใจ

จากข้อมูลข้างต้นสามารถสรุปได้ว่า ความพึงพอใจของลูกค้าเป็นความรู้สึกของลูกค้าที่ได้จากการเปรียบเทียบระหว่างการใช้อินค้าหรือบริการกับความคาดหวัง โดยความพึงพอใจและความภักดีของลูกค้ามีความสัมพันธ์กันในทิศทางเดียวกัน กล่าวคือ หากองค์กรสามารถทำให้ลูกค้าเกิดความพึงพอใจในแต่ละครั้งที่ใช้บริการได้อย่างต่อเนื่อง ความภักดีของลูกค้าจะค่อย ๆ เกิดขึ้น

ทั้งนี้ในแง่มุมมองของการท่องเที่ยว ความพึงพอใจของนักท่องเที่ยวสามารถประเมินได้หลายแง่มุม เช่น ความคาดหวังในประสบการณ์ที่ได้รับของนักท่องเที่ยวจากสถานที่ท่องเที่ยวที่ไปเยือน โดยเป็นการเปรียบเทียบระหว่าง ภาพลักษณ์ สิ่งที่พบเห็น ความรู้สึก และสิ่งที่ได้รับจากสถานที่ท่องเที่ยวในอดีตและปัจจุบัน การเปรียบเทียบระหว่างสิ่งที่นักท่องเที่ยวได้รับและใช้จ่าย ไม่ว่าจะเป็นราคาผลประโยชน์ และเวลา เป็นต้น (Yoon and Uysal, 2005: 47-48)

ในภาพที่ 2.6 แสดงแบบจำลองอิทธิพลของแรงจูงใจต่อความพึงพอใจในการท่องเที่ยวของนักท่องเที่ยว โดยความพึงพอใจ จะเกิดจากการเปรียบเทียบแรงจูงใจที่เกิดขึ้นก่อนการตัดสินใจท่องเที่ยว ของนักท่องเที่ยวกับประสบการณ์ที่นักท่องเที่ยวได้รับจากการท่องเที่ยว และหากเกิดความพึงพอใจจะนำไปสู่การกลับมาเที่ยวอีกครั้งหรือความภักดีต่อสถานที่ท่องเที่ยว (Destination Loyalty)



ภาพที่ 2.6 อิทธิพลของแรงจูงใจต่อความพึงพอใจในการท่องเที่ยว

2.5.2 งานวิจัยที่เกี่ยวข้องกับความพึงพอใจของนักท่องเที่ยว

ในการวิจัยครั้งนี้ ได้มุ่งเน้นเพื่อพัฒนาระบบสกัดความรู้แบบอัตโนมัติจากบทวิจารณ์ของลูกค้าเพื่อธุรกิจท่องเที่ยวออนไลน์ ดังนั้นการสำรวจงานวิจัยจึงเน้นที่ความพึงพอใจของลูกค้าที่เป็นนักท่องเที่ยวเป็นหลัก ซึ่งมีงานวิจัยที่เกี่ยวข้อง ดังนี้

ในปี 1978 พิแชม นิวแมนน์ และเรเชล (Pizam, Neumann and Reichel, 1978) ได้ทำการสำรวจความพึงพอใจของนักท่องเที่ยวที่แคปโคด (Cap Cod) รัฐแมสซาชูเซตส์ โดยจำแนกความพึงพอใจของนักท่องเที่ยวออกเป็น 8 ปัจจัย ได้แก่ ชายหาดที่เหมาะสม ราคา สิ่งอำนวยความสะดวกในเรื่องอาหารและเครื่องดื่ม สิ่งอำนวยความสะดวกด้านที่พัก สิ่งแวดล้อม รวมถึงย่านการค้า และทำการวิเคราะห์ข้อมูลที่ได้จากการสำรวจนักท่องเที่ยวจำนวน 685 คน การแบ่งปัจจัยในการสำรวจความพึงพอใจของนักท่องเที่ยว ส่วนใหญ่ขึ้นอยู่กับแต่ละพื้นที่ที่ทำการสำรวจและเก็บข้อมูล แต่จะคำนึงถึงปัจจัยพื้นฐานคล้ายกัน เช่น สิ่งอำนวยความสะดวก สิ่งดึงดูดความสนใจ และปัจจัยแวดล้อมอื่น ๆ ซึ่งแหล่งท่องเที่ยวที่มีลักษณะคล้ายคลึงกับแคปโคดสามารถใช้ปัจจัยเดียวกันนี้ในการสำรวจและพัฒนาพื้นที่เหล่านั้นได้

ชิ และกิว (Chi and Qu, 2008) ได้รวบรวมวิธีการศึกษาใน 3 เรื่องได้แก่ (1) การตรวจสอบทฤษฎีและหลักฐานเชิงประจักษ์ในการศึกษาโครงสร้างความสัมพันธ์ของภาพลักษณ์สถานที่ท่องเที่ยว (2) ภาพรวมความพึงพอใจของนักท่องเที่ยว และ (3) ความภักดีต่อสถานที่ท่องเที่ยว ผลการศึกษาได้สนับสนุนจุดมุ่งหมายของตัวแบบความภักดีต่อสถานที่ท่องเที่ยว ได้แก่ (1) ภาพลักษณ์สถานที่ท่องเที่ยวมีอิทธิพลต่อลักษณะความพึงพอใจของนักท่องเที่ยว (2) ภาพลักษณ์สถานที่ท่องเที่ยวและลักษณะความพึงพอใจของนักท่องเที่ยวเกิดขึ้นก่อนภาพรวมความพึงพอใจทั้งหมด และ (3) ภาพรวมและลักษณะความพึงพอใจของนักท่องเที่ยวมีผลโดยตรง และมีผลกระทบทางบวกต่อความภักดีที่มีต่อสถานที่ท่องเที่ยว

คราโคลิชิ และเนจแกมป์ (Cracolici and Nijkamp, 2008) ได้ทำการประเมินความสัมพันธ์ของการแข่งขันในการดึงดูดด้านการท่องเที่ยวจากมุมมองของนักท่องเที่ยวแต่ละบุคคล

ทางตอนใต้ของประเทศอิตาลี โดยใช้แบบสอบถามในการประเมินคุณภาพสิ่งอำนวยความสะดวกต่าง ๆ ที่มีอยู่ในพื้นที่นั้น ซึ่งข้อมูลที่ได้จากการสำรวจจะนำไปใช้ปรับปรุงนโยบายการท่องเที่ยวในเขตภูมิภาค จากผลการสำรวจทำให้ทราบถึงจุดแข็ง จุดอ่อน โอกาส และภัยคุกคาม ของภาพลักษณ์การท่องเที่ยว ในภูมิภาค และบ่งบอกว่าศักยภาพเชิงปริมาณที่ดี ไม่ได้เชื่อมโยงกับศักยภาพเชิงคุณภาพที่ยอดเยี่ยมเสมอไป

เดอ บอสควิ และซาน มาร์ติน (Del Bosque and San Martin, 2008) ได้มุ่งศึกษา การรับรู้และความพึงพอใจด้านจิตวิทยาการท่องเที่ยว โดยใช้ตัวแบบที่อธิบายความสัมพันธ์ระหว่างตัวแปรทางจิตวิทยาของนักท่องเที่ยวที่ได้พัฒนาขึ้น งานวิจัยชิ้นนี้ได้ทำการทดลองกับนักท่องเที่ยวในประเทศสเปนจำนวน 807 คน ผลการทดลองแสดงให้เห็นว่าภาพลักษณ์ที่คิดไว้ล่วงหน้าของแหล่งท่องเที่ยว 3 แห่ง มีอิทธิพลต่อความคาดหวังและความภักดีของนักท่องเที่ยวที่มีต่อแหล่งท่องเที่ยว นอกจากนี้ภาพลักษณ์ที่คิดไว้ล่วงหน้าดังกล่าวยังมีผลกระทบของความคาดหวังและอารมณ์ความพึงพอใจของนักท่องเที่ยว ซึ่งมีอิทธิพลอย่างมีนัยสำคัญต่อพฤติกรรม

ฮสุ ทสึ และวู (Hsu, Tsai and Wu, 2009) ได้วิเคราะห์ปัจจัยที่มีอิทธิพลต่อการเลือกมาท่องเที่ยวในประเทศไทยได้ทุกวัน และทำการประเมินความพึงพอใจของนักท่องเที่ยวที่มีต่อสถานที่ท่องเที่ยว โดยนำตัวแบบวิเคราะห์กระบวนการเชิงโครงสร้าง 4 ระดับมาทดสอบข้อมูลที่รวบรวมได้ ผลจากการศึกษาพบว่า 2 ปัจจัยสำคัญที่นักท่องเที่ยวเลือกมาเที่ยวประเทศไทยได้ทุกวัน คือ การมาเยี่ยมเพื่อนหรือเครือญาติ และความปลอดภัย ในส่วนของราคาพบว่ามีความสำคัญน้อยที่สุด และตึกไทเป 101 เป็นสถานที่ที่นักท่องเที่ยวให้ความสำคัญเป็นอันดับแรก

สำหรับในงานวิจัยนี้ ได้พัฒนาระบบสกัดและค้นคืนความรู้แบบอัตโนมัติจากบทวิจารณ์ของนักท่องเที่ยวเพื่อธุรกิจท่องเที่ยวออนไลน์ขึ้นมา เพื่อให้ผู้ประกอบการสามารถนำความรู้เกี่ยวกับความพึงพอใจของนักท่องเที่ยวที่ได้ ไปใช้ในการวิเคราะห์และปรับปรุงธุรกิจให้ตรงกับความต้องการของนักท่องเที่ยวมากที่สุด นอกจากนี้ ความพึงพอใจของนักท่องเที่ยวเก่า ยังช่วยสนับสนุนการตัดสินใจของนักท่องเที่ยวใหม่ในการตัดสินใจเลือกซื้อสินค้าหรือบริการต่าง ๆ ตามคำแนะนำที่ได้จากนักท่องเที่ยวด้วยกันเอง

2.6 แนวคิดเกี่ยวข้องบทวิจารณ์ออนไลน์ของนักท่องเที่ยว

ในการวิจัยเรื่องการพัฒนาาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทยนี้ ได้นำบทวิจารณ์ออนไลน์ของนักท่องเที่ยวมาใช้เป็นกรณีศึกษาในการสกัดและค้นคืนความรู้ ดังนั้นจึงได้ศึกษางานวิจัยที่เกี่ยวข้องกับบทวิจารณ์ออนไลน์ของนักท่องเที่ยวเพื่อนำมาใช้เป็นแนวทางในการวิจัย โดยจากการศึกษาพบว่าการนำบทวิจารณ์ออนไลน์มาประยุกต์ใช้ในงานทางด้านต่าง ๆ ด้วยเทคนิคที่แตกต่างกัน ดังรายละเอียดต่อไปนี้

2.6.1 ความสำคัญของบทวิจารณ์ออนไลน์

บทวิจารณ์ออนไลน์ในบริบทของงานวิจัยนี้ หมายถึง ประสบการณ์หรือความคิดเห็นของนักท่องเที่ยวที่เคยไปพักที่สถานที่นั้นมาแล้ว และนำมาวิจารณ์หรือเผยแพร่บนอินเทอร์เน็ต โดยบทวิจารณ์อาจถูกเรียกได้ในศัพท์อื่น เช่น ความคิดเห็นของนักท่องเที่ยว คำวิจารณ์ ข้อเสนอแนะต่าง ๆ จากนักท่องเที่ยว เป็นต้น จากความหมายดังกล่าวข้างต้นจะเห็นได้ว่าบทวิจารณ์ออนไลน์ของนักท่องเที่ยวเป็นข้อมูลที่แสดงถึงความพึงพอใจและไม่พึงพอใจของนักท่องเที่ยวที่มีต่อสินค้าหรือบริการของสถานที่พัก ซึ่งความพึงพอใจของนักท่องเที่ยวมีความสำคัญอย่างยิ่งต่อองค์กรตั้งที่ได้กล่าวมาแล้วในแนวคิดเกี่ยวกับความพึงพอใจของนักท่องเที่ยว นอกจากนี้ ความพึงพอใจของนักท่องเที่ยวเก่า ยังมีผลต่อการตัดสินใจเลือกซื้อสินค้าและบริการของนักท่องเที่ยวใหม่อีกด้วย นอกจากนี้ ในปัจจุบันนักท่องเที่ยวส่วนใหญ่จะใช้อินเทอร์เน็ตในการสืบค้นข้อมูลเพื่อนำมาใช้ประกอบการตัดสินใจเลือกซื้อหรือใช้สินค้าและบริการให้ได้ตรงกับความต้องการของตน ซึ่งวิธีที่นิยมใช้กันโดยทั่วไปคือ การตัดสินใจโดยใช้ข้อมูลจากบทวิจารณ์ออนไลน์ของลูกค้าที่เคยใช้สินค้าหรือบริการนั้นมาแล้ว โดยการสืบค้นจากเว็บไซต์ที่เกี่ยวข้องกับสินค้าหรือบริการนั้นและอนุญาตให้มีการเขียนบทวิจารณ์เกี่ยวกับสินค้าหรือบริการเหล่านั้นได้ เช่น เว็บไซต์ด้านการท่องเที่ยว TripAdvisor.com ซึ่งเว็บไซต์นี้นอกจากจะอนุญาตให้ผู้ใช้เขียนบทวิจารณ์แล้ว ยังมีการจัดอันดับสินค้าหรือบริการจากคะแนนที่ผู้ใช้เขียนบทวิจารณ์ได้ประเมินไว้ เพื่อความสะดวกในการใช้ข้อมูลประกอบการตัดสินใจเลือกซื้อของผู้ใช้คนอื่น ๆ ต่อไป ดังแสดงในภาพที่ 2.6



The screenshot displays the TripAdvisor page for Chatrium Suites Bangkok. The page includes a navigation menu, a search bar, and a sidebar with various travel options. The main content area features the hotel's name, address, and a 'Check Rates and Availability' section. A 'Reviews you can trust' section shows a breakdown of reviews by trip type and rating. A featured review by SandyTO is highlighted, and a 'Best deals' section lists various offers for the hotel.

ภาพที่ 2.6 เว็บไซต์ด้านการท่องเที่ยว TripAdvisor.com

จากข้อมูลข้างต้น จะเห็นได้ว่าข้อมูลบนเว็บไซต์ยังไม่ละเอียดเพียงพอสำหรับความต้องการของนักท่องเที่ยว เพราะหากนักท่องเที่ยวคนใดต้องการรายละเอียดเพิ่มเติม เช่น ต้องการทราบว่าอาหารเช้าเป็นอย่างไร ห้องพักสะอาดหรือไม่ หรือทำเลเหมาะสมหรือไม่ จะต้องตามอ่านในทุกอย่างความคิดเห็นเอง แล้วจึงวิเคราะห์หรือตัดสินใจว่าแหล่งท่องเที่ยวนั้นควรไปหรือไม่ ดังนั้น งานวิจัยนี้จึงได้มุ่งพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทยขึ้นมา เพื่อใช้เป็นเครื่องมือที่ช่วยเพิ่มความสามารถในการสกัดเอาความรู้ที่เป็นประโยชน์ต่อลูกค้า และนำความรู้ในเรื่องที่สกัดได้ไปใช้ประกอบการตัดสินใจได้อย่างสะดวกและรวดเร็วยิ่งขึ้น

2.5.2 งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์บทวิจารณ์ออนไลน์

เทอร์นี่ (Turney, 2002) ได้นำเสนอขั้นตอนวิธีอย่างง่ายในการแยกประเภทบทวิจารณ์ออนไลน์ว่ามีความหมายเชิงแนะนำหรือไม่ โดยพยากรณ์จากการคำนวณค่าเฉลี่ยของกลุ่มคำคุณศัพท์

หรือคำกริยาวิเศษณ์ บทวิจารณ์ที่ถูกจัดประเภทแล้วว่ามีความหมายในเชิงแนะนำจะมีค่าเฉลี่ยเป็นบวก
ขั้นตอนวิธีนี้มีค่าความถูกต้อง 74% โดยประเมินจากบทวิจารณ์ในเว็บไซต์ Epinions.com

เดฟ ลอว์เรนซ์ และเพนนอคค์ (Dave, Lawrence and Pennock, 2003) ได้ทำการ
จำแนกคุณสมบัติที่โดดเด่นของผลิตภัณฑ์และพัฒนาวิธีการจำแนกบทวิจารณ์ออนไลน์ในแง่บวกและแ
ลบแบบอัตโนมัติ ในส่วนของการสกัดรูปลักษณะของผลิตภัณฑ์และสกัดระดับคะแนน ได้นำเทคนิคด้าน
การสืบค้นสารสนเทศมาใช้ ผลที่ได้จากการวัดค่าที่หลากหลายและความแปรปรวนของขั้นตอนวิธี
ประเภทศึกษาสำนึก (Heuristics) ขึ้นอยู่กับสภาพแวดล้อมในการทดสอบ แต่วิธีการที่ดีที่สุดนั้นทำงาน
ได้ดีเทียบเท่าหรือดีกว่ากระบวนการเรียนรู้ของคอมพิวเตอร์แบบดั้งเดิม เมื่อวิธีการที่ดีที่สุดนี้ทำงานกับ
ประโยคเฉพาะที่ได้จากการค้นหาทางเว็บไซต์ ประสิทธิภาพจะถูกจำกัดโดยสิ่งรบกวนและความกำกวม
แต่ในบริบทของเครื่องมือทางเว็บไซต์ที่สมบูรณ์และการใช้วิธีง่าย ๆ เพื่อจัดกลุ่มประโยคให้อยู่ใน
รูปแบบคุณลักษณะ ผลที่ได้เป็นเชิงคุณภาพซึ่งสามารถนำไปใช้ประโยชน์ได้ดี

ฮู และเลีย (Hu and Liu, 2004) ได้เสนอชุดเทคนิคสำหรับการค้นหาแบบและ
ความสัมพันธ์ที่ซ่อนอยู่และการสรุปสาระของบทวิจารณ์ของผลิตภัณฑ์ โดยใช้วิธีการทำเหมืองข้อมูล
และการประมวลผลภาษาธรรมชาติ โดยมีวัตถุประสงค์ของการวิจัยเพื่อจัดเตรียมการสรุปสาระของบท
วิจารณ์ผลิตภัณฑ์จากลูกค้าที่มีจำนวนมากบนอินเทอร์เน็ต ขั้นตอนการวิจัยประกอบด้วย (1) ค้นหา
ผลิตภัณฑ์ที่ลูกค้าวิจารณ์เป็นจำนวนมาก (2) ระบุรูปแบบประโยคที่แสดงถึงทัศนคติและการตัดสินใจใน
แต่ละบทวิจารณ์ ทั้งด้านบวกและด้านลบ (3) และสรุปผล ซึ่งผลการทดลองแสดงให้เห็นว่าเทคนิคที่
นำเสนอมีผลอย่างมากในการทำให้งานวิจัยประสบความสำเร็จ และนอกจากการสรุปบทวิจารณ์
ออนไลน์จะมีประโยชน์ต่อลูกค้าแล้ว ยังมีความสำคัญมากต่อผู้ผลิตอีกด้วย

ทาโบดา และกรีฟ (Taboada and Grieve, 2004) ได้นำเสนอวิธีการจำแนกข้อความ
แบบอัตโนมัติโดยใช้กับเนื้อหาการแสดงความคิดเห็นส่วนตัว เช่น บทวิจารณ์ภาพยนตร์ บทวิจารณ์
รถยนต์ บทวิจารณ์หนังสือ โดยนำวิธีมาตรฐานในการคำนวณความหมายจากตำแหน่งของคำคุณศัพท์
ในข้อความมาประยุกต์ใช้ จากการศึกษาพบว่าประสิทธิภาพของระบบขึ้นอยู่กับประเภทของบทวิจารณ์
ที่กำลังพิจารณา นอกจากนี้ ผู้วิจัยยังได้นำเทคนิคการจำแนกภาษา (Appraisal) มาประยุกต์ใช้ในการ
สกัดคุณค่าของแต่ละบทวิจารณ์ ซึ่งก็ให้ผลลัพธ์ที่แตกต่างตามประเภทของบทวิจารณ์เช่นกัน ทั้งนี้เพื่อ
ค้นหาว่าวิธีการที่พัฒนาขึ้นมานั้น สามารถนำไปใช้ประโยชน์ในการจำแนกความแตกต่างระหว่าง
ข้อความแสดงความคิดเห็นส่วนตัวชนิดอื่น ๆ ได้หรือไม่

ไวท์แมน และเอลลิส (Whitman and Ellis, 2004) ได้ทำการวิเคราะห์บทวิจารณ์
ทางด้านดนตรี และพัฒนาระบบที่มีกลไกในการกำจัดอคติและการบรรยายที่ไม่เกี่ยวข้องกับดนตรีของผู้
วิจารณ์ โดยการเรียนรู้จากบทวิจารณ์จำนวนมาก

เซง และยี (Zheng and Ye, 2009) ได้ทำการทดลองกับบทวิจารณ์ของนักท่องเที่ยวใน
ภาษาจีนบนเว็บไซต์ ctrip.com โดยนำขั้นตอนวิธีมาให้คำแนะนำเพื่อการเรียนรู้ของเครื่อง

คอมพิวเตอรื (SVM) มาจำแนกทัศนคติของบทวิจารณ์ภาษาจีนจำนวน 40 ที่พักในเมืองเทียนจิน (Tianjin) และชองชิ่ง (Chongqing) ผลการศึกษาพบว่า (1) ขั้นตอนวิธีในการจำแนกทัศนคติของบทวิจารณ์นี้สามารถใช้ได้ดีกับบทวิจารณ์ในภาษาจีน (2) บทวิจารณ์ของนักท่องเที่ยวบนอินเทอร์เน็ตส่งผลกระทบต่ออาการจ้องที่พักในประเทศจีน (3) และผลที่ได้จากงานวิจัยนี้สามารถนำไปใช้เพิ่มศักยภาพทางธุรกิจที่พักในประเทศจีนได้อย่างไร้ข้อสงสัย

โนรูซี และโฟพูไฮ (Noroozi and Fotouhi, 2010) ได้ศึกษาอิทธิพลของเว็บไซต์เชิงความหมายที่มีต่อการตัดสินใจของผู้บริโภคในอุตสาหกรรมการท่องเที่ยว โดยมุ่งเน้นการเก็บข้อมูลจากระบบรับบทวิจารณ์สินค้าหรือบริการ ผลจากการศึกษาพบว่าผู้บริโภคที่มีความเกี่ยวข้องกับสินค้าหรือบริการต่ำจะวิเคราะห์ข้อมูลและตัดสินใจอย่างระมัดระวัง หรือไม่เชื่อถือบทวิจารณ์เลย ในขณะที่ผู้บริโภคที่มีความเกี่ยวข้องกับสินค้าหรือบริการสูงจะให้ความสำคัญกับบทวิจารณ์ และสิ่งที่เกี่ยวข้องด้วย อาทิ จำนวนผู้วิจารณ์ งานวิจัยนี้แสดงให้เห็นว่า นอกจากการอ่านเนื้อหาบทวิจารณ์หลักแล้ว บทบาทของปัจจัยทางสังคมยังมีอิทธิพลอย่างมีนัยสำคัญต่อทัศนคติของผู้บริโภคที่มีต่อบทวิจารณ์บนเว็บไซต์

จากการทบทวนวรรณกรรมดังที่กล่าวมาข้างต้น จะพบว่า งานวิจัยที่เกี่ยวข้องกับบทวิจารณ์ออนไลน์ส่วนใหญ่เป็นการจำแนกบทวิจารณ์ให้อยู่ในหมวดหมู่ที่ต้องการแบบอัตโนมัติเท่านั้น ไม่ใช่การสกัดและค้นคืนความรู้จากบทวิจารณ์เหล่านั้นแต่อย่างใด ยกเว้นงานวิจัยของ ฮู และเลีย (Hu and Liu, 2004) ที่ได้ทำการสกัดและค้นคืนความรู้ โดยการสรุปสาระสำคัญของบทวิจารณ์และนำเสนอในทั้งด้านบวกและลบให้กับลูกค้า แต่อย่างไรก็ตาม ลูกค้าไม่สามารถทราบถึงจำนวนคำวิจารณ์ต่าง ๆ อย่างละเอียด ว่ามีผู้ที่แสดงความคิดเห็นทั้งด้านบวกหรือด้านลบในแต่ละเรื่องจำนวนเท่าใด ซึ่งในการจัดลำดับความสำคัญนั้น จำนวนของการแสดงความคิดเห็นนับว่ามีผลอย่างมากต่อการจัดลำดับความสำคัญ หรือการตัดสินใจเลือกสินค้าและบริการของลูกค้า (Saaty, 2006: 4-5) ดังนั้น งานวิจัยนี้จึงนำเสนอระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทยอย่างละเอียด มีการแสดงจำนวนความคิดเห็นของแต่ละประโยคคำวิจารณ์ โดยการนำเอาเทคนิคต่าง ๆ ที่มีอยู่ในปัจจุบันเข้ามาประยุกต์ใช้ เช่น ออนโทโลยี มาใช้เป็นฐานความรู้ในการวิเคราะห์คำศัพท์ การสกัด จัดเก็บ และค้นคืนความรู้ นอกจากนี้ยังนำตัวแปลภาษามาช่วยในการวิเคราะห์เชิงโครงสร้าง รวมทั้งพัฒนาวิธีการคำนวณระดับคะแนนความพึงพอใจขึ้นมาใช้ เพื่อให้ได้มาซึ่งผลลัพธ์ที่มีความถูกต้องมากที่สุด

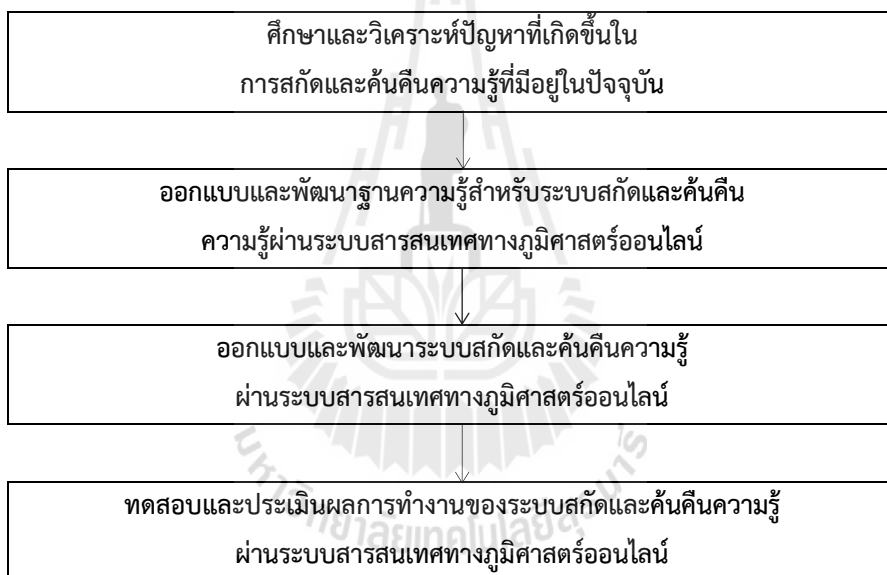
บทที่ 3

วิธีดำเนินการวิจัย

ในบทนี้กล่าวถึง วิธีการวิจัย เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล โดยมีรายละเอียด ดังนี้

3.1 วิธีการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงประยุกต์ (Applied Research) โดยประยุกต์ใช้แนวทางของ วงจรการพัฒนาาระบบ (System Development Life Cycle: SDLC) ในการพัฒนาระบบสกัดและคั่น คั้นความรู้แบบอัตโนมัติผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ โดยมีขั้นตอนดำเนินการวิจัยดังนี้



ภาพที่ 3.1 วิธีการดำเนินงานวิจัย

3.1.1 ศึกษาและวิเคราะห์ปัญหาที่เกิดขึ้นในการสกัดและคั่นคั้นความรู้ที่มีอยู่ในปัจจุบัน

จากการทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้องกับการสกัดและคั่นคั้นความรู้ในบทที่ 2 สามารถสรุปเป็นวิธีการ การใช้งาน และสารสนเทศหรือความรู้ที่สกัดได้ พร้อมทั้งกำหนดรายละเอียดต่าง ๆ ของปัญหาได้ชัดเจน และสามารถกำหนดปัจจัยที่ใช้ในระบบสกัดความรู้แบบอัตโนมัติได้ โดยมีทั้งปัจจัยที่ผู้ใช้ต้องกำหนดเพื่อป้อนเข้าสู่ระบบ และปัจจัยที่ระบบกำหนดขึ้น โดยปัจจัยทั้งสองนั้นถูกนำมาประมวลผลเพื่อให้ได้ผลลัพธ์หรือสารสนเทศตามวัตถุประสงค์ โดยสามารถ ปัจจัยนำเข้าและผลลัพธ์ได้ดังตารางที่ 3.1

ตารางที่ 3.1 ปัจจัยที่ใช้ในระบบสกัดและค้นคืนความรู้แบบอัตโนมัติและผลลัพธ์ที่ได้

ปัจจัยที่เกี่ยวข้อง	ผลลัพธ์ที่ได้
<p><u>ปัจจัยที่ผู้ใช้งานกำหนด (User-defined Variables)</u></p> <ul style="list-style-type: none"> - บทความที่มีต่อสถานที่ที่เคยไปมาแล้วนั้น หรือ - สถานที่ที่ต้องการสืบค้น 	<ul style="list-style-type: none"> - อันดับของคุณลักษณะต่าง ๆ (Features) ที่เกี่ยวข้องกับสถานที่นั้น ๆ อย่างละเอียด - ความรู้ที่เกี่ยวข้องกับสถานที่ที่ต้องการสืบค้นผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์
<p><u>ปัจจัยที่ระบบกำหนด (Pre-defined Variables)</u></p> <ul style="list-style-type: none"> - ออนโทโลยีด้านการท่องเที่ยว (Ontology) - คลังคำศัพท์เฉพาะ (Terminology) - คลังคำศัพท์ทั่วไป (Dictionary) - กฎไวยากรณ์ไม่พึ่งบริบท (Context Free Grammar) - พิกัดและชื่อของสถานที่พักและสิ่งอำนวยความสะดวกรอบ ๆ ที่พักนั้น 	

3.1.2 ออกแบบและพัฒนาฐานความรู้

ฐานความรู้ที่ออกแบบและพัฒนาขึ้นมาี้ ถูกนำไปใช้เพื่อพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ ซึ่งประกอบด้วยออนโทโลยีด้านการท่องเที่ยว คลังคำศัพท์เฉพาะที่ใช้สำหรับระบบนี้ คลังคำศัพท์ทั่วไป กฎไวยากรณ์ไม่พึ่งบริบท และพิกัดและชื่อของสถานที่พักและสิ่งอำนวยความสะดวกรอบ ๆ ที่พักนั้น โดยมีรายละเอียดดังต่อไปนี้

3.1.2.1 ออนโทโลยีด้านการท่องเที่ยว

งานวิจัยนี้ได้ทำการพัฒนาออนโทโลยีด้านการท่องเที่ยว (Tourism Ontology) โดยปรับปรุงจากออนโทโลยีด้านการท่องเที่ยวออนไลน์ (Class Hierarchy for the e-Tourism Ontology Version 8) ของเสียร์เพล์ พรานท์เนอร์ และบาชเลชเนอร์ (Siorpaes, Prantner and Bachlechner, 2004) ซึ่งประกอบด้วยโครงสร้างของคลาส (Class) จำนวน 9 คลาส และคุณสมบัติ (Properties) ของแต่ละคลาส รวมทั้งสิ้น 119 คุณสมบัติ ซึ่งในจำนวนนี้เป็นคุณสมบัติเฉพาะของคลาสที่พักจำนวน 51 คุณสมบัติ ดังแสดงโครงสร้างของคลาสในภาพที่ 3.2

- 1) Accommodation (ที่พัก)
- 2) Activity (กิจกรรม)
- 3) ContactData (ข้อมูลการติดต่อ)
- 4) DateTime (วัน เวลา เปิดทำการ)
 - OpeningHours (ช่วงวัน เวลา เปิดทำการ)
 - Period (ระยะเวลา)
 - DatePeriod (วัน)
 - TimePeriod (เวลา)
 - Season (ฤดูกาล)
- 5) Event (กิจกรรมพิเศษ)
- 6) Infrastructure (โครงสร้างพื้นฐาน)
- 7) Location (ตำแหน่งที่ตั้ง)
 - GPSCoordinates (พิกัด)
 - PostalAddress (ที่อยู่)
- 8) Room (ห้อง)
 - ConferenceRoom (ห้องประชุม)
 - Guestroom (ห้องพัก)
- 9) Ticket (ตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว)

ภาพที่ 3.2 โครงสร้างคลาสของออนโทโลยีด้านการท่องเที่ยวของเสียร์เพล์ พรานท์เนอร์ และบาชเลชเนอร์ (Siorpaes, Prantner and Bachlechner, 2004)

อย่างไรก็ตาม ออนโทโลยีด้านการท่องเที่ยวดังกล่าวยังไม่สามารถรองรับระบบสกัดและค้นคืนความรู้ที่พัฒนาขึ้นมา ซึ่งนั่นในการวิจัยนี้จึงได้ปรับปรุงออนโทโลยีนั้น โดยเฉพาะในเรื่องของคุณสมบัติของคลาสที่พัก โดยการนำบทวิจารณ์ของนักท่องเที่ยวจำนวน 400 บทวิจารณ์มาวิเคราะห์เพื่อหาคุณสมบัติหรือคุณลักษณะที่อยู่ในขอบเขตเรื่องที่พัก (Accommodation Feature) โดยการคัดเลือกคำในบทวิจารณ์จากวิธีการวัดค่าน้ำหนักของคำตามขั้นตอนวิธีของรอกชิโอ (Rocchio: TF-IDF Weighting) ดังสูตรคำนวณในสมการที่ 3-1 (Salton and Buckley, 1988: 513-523)

$$W_d = f_{w,d} \times \log(|D|/f_{w,D}) \quad (3-1)$$

โดย W_d คือ ค่าน้ำหนักของคำที่ปรากฏในเอกสาร (d)

$f_{w,d}$ คือ ความถี่หรือจำนวนครั้งที่ปรากฏคำ (w) ในเอกสาร (d)

$|D|$ คือ จำนวนของเอกสารทั้งหมด

$f_{w,D}$ คือ จำนวนของเอกสารทั้งหมดที่มีคำ (w) ปรากฏอยู่

ทั้งนี้ ข้อมูลบทวิจารณ์ทั้งหมด 400 บทวิจารณ์ที่นำมาวิเคราะห์สามารถจำแนกตามระดับคะแนนที่มีผู้วิจารณ์ได้ประเมินไว้ โดยแบ่งเป็น 5 กลุ่ม ๆ ละ 80 บทวิจารณ์ จากนั้นจึงนำคุณสมบัติที่วิเคราะห์ได้มาเพิ่มเติมลงในออนโทโลยีด้านการท่องเที่ยว ทั้งนี้ โครงสร้างของคลาสของออนโทโลยีด้านการท่องเที่ยวที่ได้ปรับปรุงขึ้นมาใหม่แสดงในภาพที่ 3.3

<p>1) Accommodation (ที่พัก)</p> <ul style="list-style-type: none"> ▪ Accommodations (ที่พัก) <ul style="list-style-type: none"> ▪ HomeStay (โฮมสเตย์) ▪ Hotel (ที่พัก) ▪ Park (วนอุทยาน หรืออุทยานแห่งชาติ) ▪ Resort (รีสอร์ท) ▪ Bathroom (ห้องน้ำ) ▪ Room (ห้อง) <ul style="list-style-type: none"> ▪ ConferenceRoom (ห้องประชุม) ▪ Guestroom (ห้องพัก) <p>2) Activity (กิจกรรม)</p> <ul style="list-style-type: none"> ▪ Adventure (ผจญภัย) ▪ Relaxation (พักผ่อน) ▪ Sightseeing (ชมสถานที่ หรือดูสัตว์) ▪ Sport (กีฬา) <p>3) Amenity (สถานที่อำนวยความสะดวก)</p> <ul style="list-style-type: none"> ▪ CarServices (บริการเกี่ยวกับรถยนต์ เช่น ปั่นน้ำมัน อู่ซ่อมรถ) ▪ EmergencyServices (บริการเหตุฉุกเฉิน เช่น โรงพยาบาล สถานีตำรวจ) ▪ Resturant (ร้านอาหาร) ▪ ServiceShops (ร้านค้า หรือห้างสรรพสินค้า) <p>4) Attraction (สถานที่ท่องเที่ยว)</p> <ul style="list-style-type: none"> ▪ AgroTourism (ท่องเที่ยวเชิงเกษตร) ▪ CulturalTourism (ท่องเที่ยวทางวัฒนธรรม) ▪ HealthTourism (ท่องเที่ยวเพื่อสุขภาพ) ▪ HistoricalTourism (ท่องเที่ยวทางประวัติศาสตร์) ▪ NaturalTourism (ท่องเที่ยวธรรมชาติ) <p>5) ContactData (ข้อมูลการติดต่อ)</p> <p>6) DateTime (วัน เวลา เปิดทำการ)</p> <p>7) Event (กิจกรรมพิเศษ)</p> <p>8) Location (ตำแหน่งที่ตั้ง)</p> <ul style="list-style-type: none"> ▪ GPSCoordinates (พิกัด) ▪ PostalAddress (ที่อยู่) <p>9) Rating (การจัดอันดับ)</p> <p>10) Ticket (ตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว)</p> <p>11) Skos: concept (คลังคำเหมือน)</p>
--

ภาพที่ 3.3 โครงสร้างคลาสของออนโทโลยีด้านการท่องเที่ยวที่สร้างขึ้นในงานวิจัยนี้

ซึ่งออนโทโลยีที่พัฒนาขึ้นมาี้ ประกอบด้วยคลาสหลักจำนวน 10 คลาส โดยได้ทำการปรับปรุงออนโทโลยีด้านการท่องเที่ยว ดังนี้

1. เพิ่มเติมคลาส ได้แก่ คลาสย่อยของคลาสที่พัก คลาสย่อยของกิจกรรม คลาสย่อยของสถานที่อำนวยความสะดวก คลาสย่อยของสถานที่ท่องเที่ยว และคลาสการจัดอันดับ
2. ลดคลาส ได้แก่ คลาสย่อยของวัน เวลา เปิดทำการ เนื่องจากคลาสย่อยดังกล่าวออกแบบไม่เหมาะสมต่อการนำไปใช้จริง จึงมีการเปลี่ยนโครงสร้างของการเก็บข้อมูลเป็นแบบคุณสมบัติแทนที่การสร้างคลาสย่อยดังกล่าว
3. ย้ายคลาส ได้แก่ คลาสห้อง ไปเป็นคลาสย่อยของคลาสที่พัก
4. เพิ่มเติมคุณสมบัติของคลาสที่พัก โดยได้เพิ่มคุณสมบัติใหม่ที่ได้จากการวิเคราะห์บทวิจารณ์มาทั้งสิ้นจำนวน 76 คุณสมบัติ

นอกจากนี้ยังมีการพัฒนาคลังคำเหมือน ซึ่งสร้างขึ้นโดยใช้สคอส (Simple Knowledge Organization System: SKOS) หรือเครื่องมือเสริมที่ช่วยในการสร้างคลังคำที่มีความหมายเหมือนกัน แต่เขียนต่างกัน เพื่อให้ง่ายแก่การสกัดและค้นคืนข้อมูล ในการสร้างสคอสนั้นจะทำการกำหนดคำที่เป็นตัวแทนของกลุ่มคำเหมือนแต่ละคำ เพื่อให้คำเหล่านี้มีการจัดหมวดหมู่ในรูปแบบเดียวกัน ตัวอย่างเช่น คำว่า “เครื่องปรับอากาศ หรือ Air Conditioner” นักท่องเที่ยวแต่ละคนอาจใช้คำที่แตกต่างกัน เช่น a.c., a/c, air condition ในออนโทโลยีจะใช้คำที่เป็นตัวแทนคือ AirCondition

ทั้งนี้ โครงสร้างของออนโทโลยีด้านการท่องเที่ยวที่ได้พัฒนาขึ้นมาั้น ได้แสดงรายละเอียดของคลาส คุณสมบัติ และคลังคำเหมือนทั้งหมดไว้ในภาคผนวก ก

3.1.2.2 คลังคำศัพท์เฉพาะ

คลังคำศัพท์เฉพาะ (Terminology) ได้มาจากการวิเคราะห์บทวิจารณ์เกี่ยวกับที่พักของนักท่องเที่ยวจำนวน 400 บทวิจารณ์เช่นเดียวกับออนโทโลยีด้านการท่องเที่ยว โดยคลังคำศัพท์เฉพาะเป็นแหล่งรวบรวมคำคุณศัพท์ (Adjective) ที่ได้กำหนดระดับคะแนนไว้ และคำกริยาวิเศษณ์ (Adverb) ที่กำหนดการเพิ่ม-ลดระดับคะแนนไว้เช่นกัน ตัวอย่างคำคุณศัพท์และระดับคะแนนที่กำหนด เช่น Effective = 4 และ Bad = 2 หรือตัวอย่างคำกริยาวิเศษณ์ที่ได้กำหนดการเพิ่ม-ลดระดับคะแนน เช่น Very = ± 1 และ Not = ± 2 ทั้งนี้ การให้ระดับคะแนนของคำได้จากการยืนยันของผู้เชี่ยวชาญด้านภาษา โดยคะแนนแต่ละระดับจะแสดงถึงความพึงพอใจที่แตกต่างกัน ดังภาพที่ 3.4 โดยระดับคะแนนที่ 1 หมายถึง แย่มาก (Terrible) ระดับคะแนนที่ 2 หมายถึง แย่ (Poor) ระดับคะแนนที่ 3 หมายถึง ปานกลาง (Average) ระดับคะแนนที่ 4 หมายถึง ดี (Good) และระดับคะแนนที่ 5 หมายถึง ดีมาก (Excellent)



ภาพที่ 3.4 คะแนนที่แสดงระดับความพึงพอใจที่แตกต่างกัน

งานวิจัยนี้ได้จำแนกคำศัพท์ในคลังคำศัพท์เฉพาะออกเป็น 3 ประเภทใหญ่ ๆ ได้แก่ (1) กลุ่มคำคุณศัพท์ (Adjectives) ซึ่งมีการกำหนดระดับคะแนนไว้ตายตัว; (2) กลุ่มคำกริยาวิเศษณ์ (Adverbs) ซึ่งกำหนดการเพิ่ม-ลดระดับคะแนนไว้ เมื่อนำคำกลุ่มนี้ไปประมวลผลจะทำให้ระดับคะแนนที่สกัดได้มีผลลัพธ์ไปในทางเดียวกัน กล่าวคือหากผู้ใช้วิจารณ์ในแง่ดี ระดับคะแนนที่สกัดได้ก็จะยิ่งเพิ่มมากขึ้น และ (3) กลุ่มคำกริยาวิเศษณ์แบบพิเศษ (Negative Adverbs) ซึ่งกำหนดการเพิ่ม-ลดระดับคะแนนไว้ เช่นเดียวกับกลุ่มคำกริยาวิเศษณ์ แต่แตกต่างกันตรงที่เมื่อนำคำกลุ่มนี้ไปประมวลผลจะทำให้ระดับคะแนนที่สกัดได้มีผลลัพธ์ไปในทางตรงกันข้าม กล่าวคือหากผู้ใช้วิจารณ์ในแง่ดี ระดับคะแนนที่สกัดได้ก็จะถูกลดทอนลง ทั้งนี้ ได้แสดงรายละเอียดคลังคำศัพท์เฉพาะเหล่านี้ไว้ในภาคผนวก ก

3.1.2.3 คลังคำศัพท์ทั่วไป

คลังคำศัพท์ทั่วไป (Dictionary) งานวิจัยนี้ได้นำข้อมูลคำศัพท์ภาษาอังกฤษทั่วไปที่ต้องใช้ในการวิเคราะห์บทวิจารณ์มาจากพจนานุกรมสี่อเล็กทรอนิกส์เล็กชิตรอนรุ่นที่ 3 (Lexitron Version 3.0 beta) ซึ่งพัฒนาโดยหน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ โดยพจนานุกรมนี้ ผู้พัฒนาได้เปิดให้บริการกับประชาชนทั่วไปแบบไม่เสียค่าใช้จ่าย เพื่อให้ผู้ที่สนใจสามารถนำไปใช้งานทั้งในด้านการศึกษาและการวิจัยทางด้านภาษาและคอมพิวเตอร์ ปัจจุบันพจนานุกรมเล็กชิตรอนประกอบด้วยคำศัพท์ภาษาอังกฤษกว่า 79,000 คำ ซึ่งได้จากการนำเทคโนโลยีทางคอมพิวเตอร์มาช่วยในการรวบรวมและคัดเลือกคำประโยค หรือข้อความที่มีใช้จริงและมีอัตราการปรากฏสูงในบริบทต่าง ๆ ของการใช้ภาษา ทั้งจากแหล่งข้อมูลและข่าวสารที่เผยแพร่ทางอินเทอร์เน็ตและแหล่งข้อมูลอื่น ๆ ที่เชื่อถือได้ เช่น วรรณกรรม บทความ ข้อมูลข่าวสารจากหนังสือพิมพ์ (งานเทคโนโลยีประมวลผลข้อความ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ, www., 2555)

3.1.2.4 กฎไวยากรณ์ไม่พึ่งบริบท

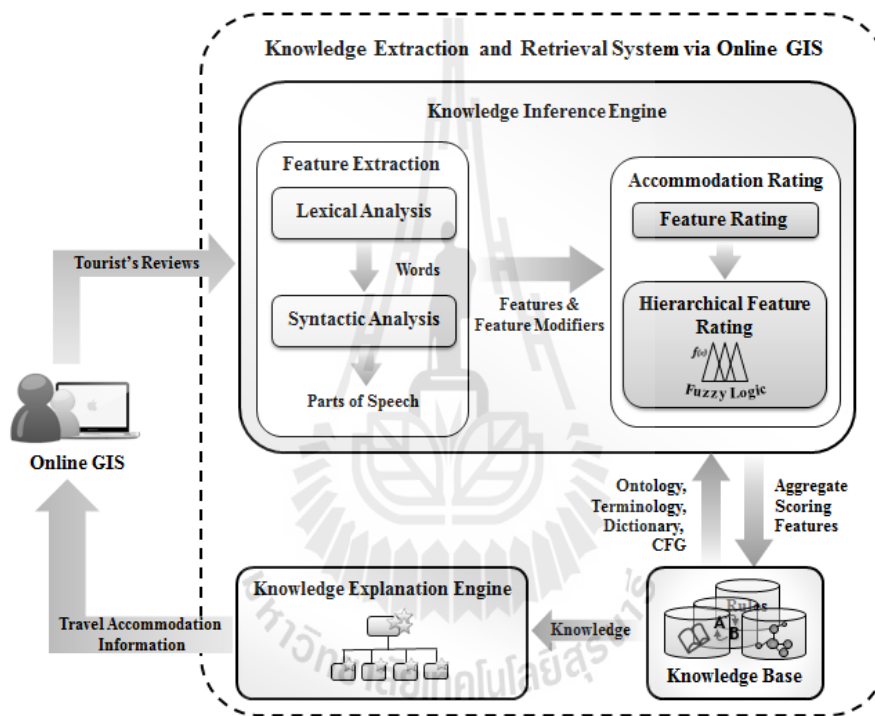
กฎไวยากรณ์ไม่พึ่งบริบท (Context Free Grammer: CFG) เป็นกลุ่มของกฎไวยากรณ์ที่เขียนขึ้นโดยแสดงเป็นสัญลักษณ์ทางภาษา เพื่อใช้สำหรับตรวจสอบไวยากรณ์ของประโยค คำวิจารณ์ในขั้นตอนการแปลภาษา กลุ่มของกฎไวยากรณ์ที่ใช้ในงานวิจัยนี้แสดงดังภาพที่ 3.5

S -> NP VP	CONJP -> CONJ VP	VP -> AUXP N
S -> N VP	CONJP -> CONJ V	AUXP -> AUX NOT
S -> NP V	CONJP -> CONJ AUXP	VP -> VB AP
S -> N V	AUXP -> AUX NP	VP -> VB ADJ
S -> AND S	AUXP -> AUX N	VP -> VBP ADJ
NP -> PRON	VP -> V	VP -> AND V
NP -> DET NP	VP -> ADV VP	VP -> V VP
NP -> DET N	VP -> ADV V	VP -> NOT AP
NP -> PRON N	VP -> VP NP	PP -> P NP
NP -> N	VP -> V NP	PP -> P N
NP -> AP N	VP -> VP N	PP -> P VP
NP -> ADJ N	VP -> V N	PP -> P V
NP -> NP PP	VP -> V PP	AP -> ADJ
NP -> N PP	VP -> VP PP	AP -> ADV AP
NP -> NP CONJP	VP-> AUX VP	AP -> ADV ADJ
NP -> N CONJP	VP -> AUXP VP	AP -> ADJ PP
NP -> AND N	VP -> AUXP AP	AP -> ADJ
NP -> N NP	VP -> AUXP NP	AP -> AND ADJ
NP -> AND VP	VP -> AUX NP	AP -> ADJ
NP -> AND AP	VP -> AUX N	
CONJP -> CONJ NP	VBP -> VB NOT	

ภาพที่ 3.5 กฎไวยากรณ์ไม่พึ่งบริบทของระบบ

3.1.3 ออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย

ระบบสกัดความรู้แบบอัตโนมัติที่พัฒนาในงานวิจัยนี้ มุ่งเน้นที่การวิเคราะห์ความหมายของบทวิจารณ์เกี่ยวกับสถานที่พักสำหรับนักท่องเที่ยว โดยผู้ใช้งานจะต้องป้อนบทวิจารณ์เกี่ยวกับสถานที่พักที่เคยไปมาแล้วเข้าสู่ระบบโดยผ่านส่วนปฏิสัมพันธ์กับผู้ใช้ (User Interface) ซึ่งมีระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ (Online GIS) จากนั้นระบบจะทำการสกัดและค้นคืนความรู้ตามสถาปัตยกรรมโครงสร้างของระบบที่แสดงดังภาพที่ 3.6 ซึ่งนำออนไลน์ที่ผ่านการปรับปรุงอย่างเหมาะสมมาประยุกต์ใช้งาน เพื่อใช้เพิ่มความสามารถในการสกัดและค้นคืนความรู้

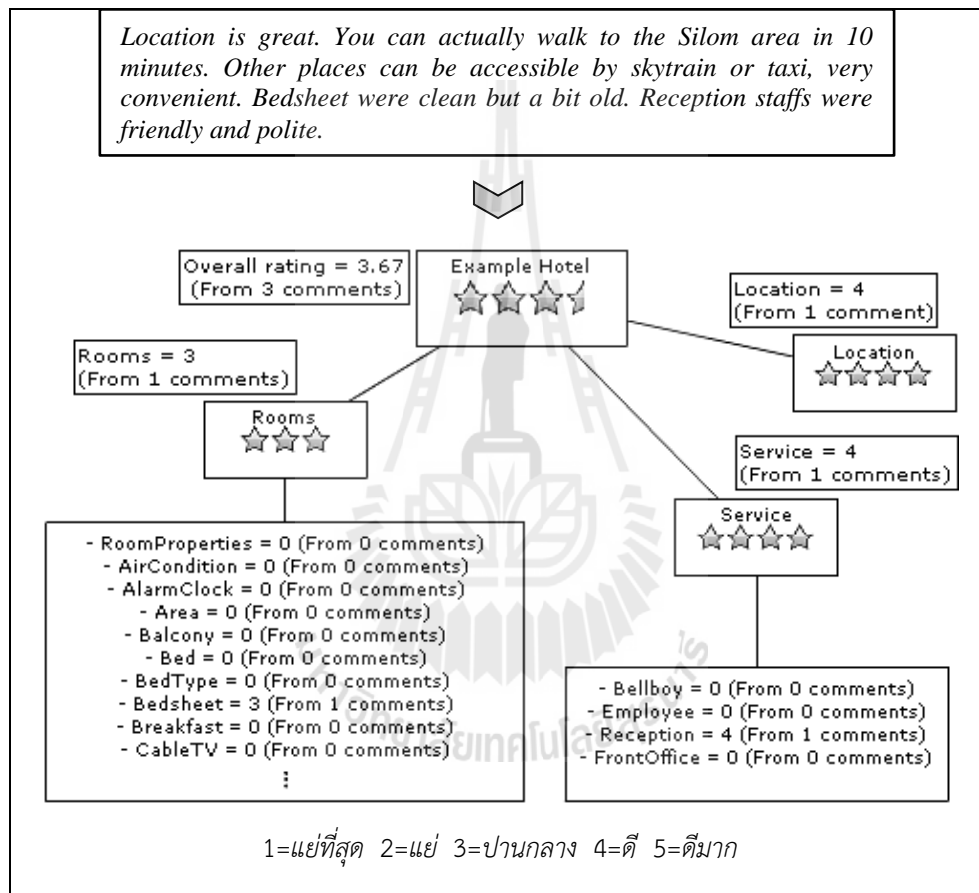


ภาพที่ 3.6 โครงสร้างของระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย

โครงสร้างการทำงานหลักของระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย ได้แก่ ส่วนติดต่อผู้ใช้ซึ่งผนวกกับระบบสารสนเทศทางภูมิศาสตร์ออนไลน์ (Online GIS) ฐานความรู้ (Knowledge Base) ส่วนอนุมานความรู้ (Knowledge Inference Engine) และส่วนอธิบายความรู้ (Knowledge Explanation Engine) โดยอธิบายรายละเอียดได้ดังต่อไปนี้

3.1.3.1 ส่วนติดต่อผู้ใช้

ส่วนติดต่อผู้ใช้เป็นส่วนโต้ตอบกับผู้ใช้หรือนักท่องเที่ยวซึ่งใช้งานระบบผ่านทางเว็บเบราว์เซอร์ ส่วนนี้ได้ถูกออกแบบและพัฒนาให้มีความเป็นมิตรต่อผู้ใช้งานโดยใช้ภาพและสัญลักษณ์กราฟิก (Graphic User Interface: GUI) ทำให้ผู้ใช้สามารถค้นหาข้อมูลของสถานที่พักสมัครสมาชิก และเข้าสู่ระบบเพื่อป้อนบทวิจารณ์เกี่ยวกับสถานที่พักได้โดยสะดวก ทั้งนี้ เมื่อระบบทำการสกัดความรู้จากบทวิจารณ์เรียบร้อยแล้ว ผลลัพธ์ที่ได้จะถูกแสดงผลผ่านทางส่วนติดต่อผู้ใช้ในรูปแบบแผนภาพโครงสร้างต้นไม้ ดังแสดงในภาพที่ 3.7 เพื่อสรุปการวิจารณ์ของสถานที่พักแห่งนั้น



ภาพที่ 3.7 ส่วนติดต่อผู้ใช้ซึ่งแสดงข้อมูลนำเข้าและผลลัพธ์ของระบบสกัดและค้นคืนความรู้

3.1.3.2 ฐานความรู้

ฐานความรู้เป็นส่วนสำคัญของระบบที่ทำให้ทราบถึงคำศัพท์ต่าง ๆ ที่มีในบทวิจารณ์เพื่อนำไปใช้ในการสกัดคุณลักษณะที่สำคัญของสถานที่พักที่กล่าวถึงในบทวิจารณ์นั้นออกมา รวมทั้งทราบถึงระดับคะแนนของคำศัพท์เฉพาะต่าง ๆ ที่นำไปใช้ในการคำนวณระดับคะแนนของสถานที่พักนั้น ๆ นอกจากฐานความรู้จะเก็บสะสมข้อมูลต่าง ๆ ดังที่กล่าวมาแล้ว ยังเก็บระดับคะแนนของสถานที่พักที่ระบบคำนวณได้จากการสกัดความรู้จากบทวิจารณ์ที่ผู้ใช้นำเข้ามาแบบเรียลไทม์ เพื่อ

เป็นระดับคะแนนของสถานที่พักนั้น ๆ โดยวิธีในการได้มาซึ่งระดับคะแนนของสถานที่พักนั้นจะกล่าวถึงในส่วนอนุมานความรู้ถัดไป

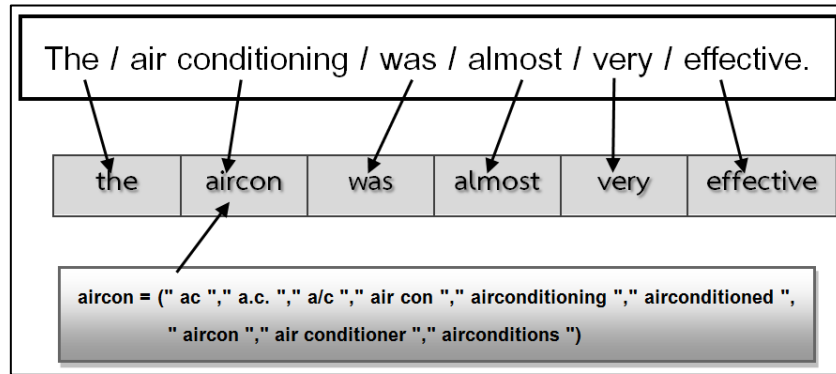
3.1.3.3 ส่วนอนุมานความรู้

ส่วนอนุมานความรู้เป็นส่วนสำคัญที่สุดของระบบ ซึ่งมีหน้าที่ในการสกัดความรู้จากบทวิจารณ์แบบอัตโนมัติ และนำความรู้ที่ได้มาเก็บใส่ฐานความรู้ เพื่อรองรับการสืบค้นจากผู้ใช้ต่อไป โดยส่วนอนุมานความรู้นี้ประกอบด้วย 2 กระบวนการหลัก ได้แก่ กระบวนการสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์ (Feature Extraction) และกระบวนการคำนวณระดับคะแนนของที่พัก (Accommodation Rating)

กระบวนการที่ 1 การสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์

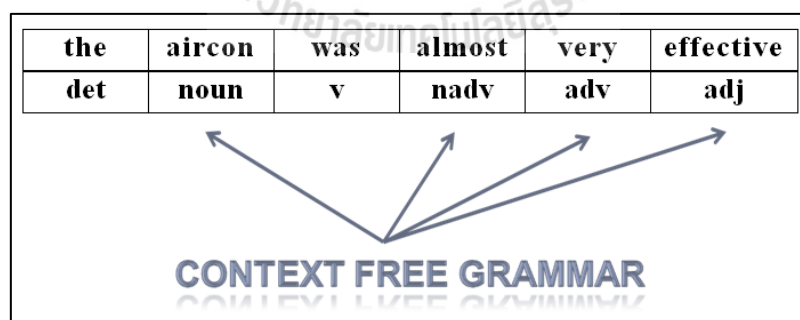
กระบวนการสกัดคุณลักษณะที่สำคัญของสถานที่พักออกจากบทวิจารณ์ออนไลน์ (Feature Extraction) สามารถแบ่งการทำงานออกเป็น 2 ส่วน คือ การวิเคราะห์คำศัพท์ (Lexical Analysis) และการวิเคราะห์เชิงโครงสร้าง (Syntactic Analysis) ซึ่งมีรายละเอียดดังต่อไปนี้

1) **การวิเคราะห์คำศัพท์ (Lexical Analysis)** ทำหน้าที่ในการสกัดคำศัพท์ต่าง ๆ ออกจากบทวิจารณ์ของนักท่องเที่ยว เพื่อจะนำคำศัพท์เหล่านั้นไปใช้ในกระบวนการวิเคราะห์เชิงโครงสร้างต่อไป โดยในการวิเคราะห์คำศัพท์ ประกอบด้วย การตัดคำ การแปลงคำศัพท์ให้อยู่ในรูปแบบที่เหมาะสม การตรวจสอบคำผิดเพื่อเปลี่ยนให้เป็นคำที่ถูกต้อง และการตรวจสอบคำที่มีความหมายเหมือนกัน แต่เขียนต่างกัน โดยในงานวิจัยนี้ได้นำบทวิจารณ์ของนักท่องเที่ยวที่เป็นภาษาอังกฤษมาเป็นข้อมูลตั้งต้นมา หลังจากนั้นได้นำข้อความในบทวิจารณ์มาตัดคำตามช่องว่าง และแปลงคำศัพท์ให้อยู่ในรูปแบบที่เหมาะสมสำหรับการสกัดความรู้ อันได้แก่ การตรวจสอบคำผิดเพื่อเปลี่ยนให้เป็นคำที่ถูกต้อง และการตรวจสอบคำที่มีความหมายเหมือนกัน แต่เขียนต่างกัน โดยเปรียบเทียบจากคลังคำเหมือนที่อยู่ในออนโทโลยี เพื่อเปลี่ยนให้คำเหล่านั้นเป็นคำเดียวกันทั้งหมด อาทิ คำว่า “air conditioning” เปลี่ยนเป็นคำว่า “aircon” ดังแสดงในภาพที่ 3.8



ภาพที่ 3.8 ตัวอย่างการทำงานของตัววิเคราะห์คำศัพท์

2) การวิเคราะห์เชิงโครงสร้าง (Syntactic Analysis) ทำหน้าที่แยกแยะส่วนของประโยค คำวิจารณ์ โดยการจำแนกว่าส่วนใดคือ ประธาน กริยา กรรม และส่วนขยายของประโยค ซึ่งในส่วนนี้ได้ นำออนโทโลยีด้านการท่องเที่ยวเข้ามาช่วยวิเคราะห์ด้วย โดยระบบจะเปรียบเทียบคำศัพท์ที่ตัดได้ใน ขั้นตอนที่แล้วกับคำศัพท์ในออนโทโลยี เพื่อค้นหาคำนามซึ่งเป็นคำสำคัญที่บ่งบอกถึงเรื่อง ที่นักท่องเที่ยวกล่าวถึงและนำมาใช้ในการวิเคราะห์ระดับคะแนนในเรื่องนั้น เมื่อทราบคำสำคัญแล้ว ระบบ จะจำแนกประเภทของคำศัพท์ทั้งหมดเฉพาะที่อยู่ในประโยคคำวิจารณ์นั้น และทำการวิเคราะห์หา ความสัมพันธ์กันระหว่างคำในประโยคตามทฤษฎีไวยากรณ์ไม่พึงบริบท ดังภาพที่ 3.9 เพื่อให้ทราบว่า คำคุณศัพท์คำใดขยายคำนามคำใด ในกรณีที่มีคำคุณศัพท์และคำนามปรากฏอยู่ในประโยคมากกว่า 1 คู่ ทั้งนี้ การวิเคราะห์หาความสัมพันธ์กันระหว่างคำของประโยคคำวิจารณ์ มีไว้เพื่อเตรียมพร้อมสำหรับ กระบวนการคำนวณระดับคะแนนของที่พักต่อไป



ภาพที่ 3.9 ตัวอย่างการทำงานของตัววิเคราะห์เชิงโครงสร้าง

กระบวนการที่ 2 กระบวนการคำนวณระดับคะแนนของที่พัก

กระบวนการนี้ทำหน้าที่ในการคำนวณระดับคะแนนคุณลักษณะของที่พัก และระดับคะแนน โดยรวมของที่พักนั้น ๆ โดยมีจุดมุ่งหมายเพื่อออกแบบวิธีการในการคำนวณระดับคะแนนให้ใกล้เคียง

กับคะแนนที่ผู้ใช้ได้ให้ไว้มากที่สุด ซึ่งวิธีการในการคำนวณสามารถแบ่งการทำงานออกเป็น 2 ส่วน คือ การคำนวณระดับคะแนนคุณลักษณะของที่พัก (Feature Rating) และการคำนวณระดับคะแนนคุณลักษณะของที่พักแบบเป็นลำดับชั้น (Hierarchical Feature Rating) เพื่อหาระดับคะแนนโดยรวมซึ่งมีรายละเอียดดังต่อไปนี้

1) การคำนวณระดับคะแนนคุณลักษณะของที่พัก (Feature Rating) ซึ่งคำว่า คุณลักษณะ (feature) หมายถึง คำนามหรือคำสำคัญซึ่งปรากฏอยู่ในแต่ละประโยค การคำนวณระดับคะแนนคุณลักษณะ (feature rating) จะหมายถึง การคำนวณระดับความพึงพอใจของแต่ละคำคุณลักษณะ โดยปรกติแล้วการคำนวณระดับคะแนนคุณลักษณะจะถูกกำหนดโดยคะแนนของคำคุณศัพท์ที่ขยายคำคุณลักษณะซึ่งจะถูกเก็บอยู่ในคลังคำศัพท์เฉพาะ ตัวอย่างเช่น ประโยคที่ว่า “The room is clean.” คำคุณลักษณะคือ คำว่า “Room” คำคุณศัพท์ที่ขยายคำคุณลักษณะคือ คำว่า “Clean” ซึ่งมีคะแนนคำคุณศัพท์เป็น 4 คะแนน แต่อย่างไรก็ตามประโยควิจารณ์ต่างมักจะอยู่ในรูปแบบของเอเอซี (AAC: Adverb Adjective Combination) (Benamara, Cesarano, Picariello, Subrahmanian, Reforgiato, 2006: 1-7) ซึ่งจะประกอบไปด้วยคำวิเศษณ์ที่ขยายคำคุณศัพท์อีกครั้ง ตัวอย่างเช่น คำว่า “very good” ซึ่งประกอบไปด้วย คำวิเศษณ์คือคำว่า “very” และคำคุณศัพท์คือคำว่า “good” ดังนั้นการคำนวณระดับคะแนนคุณลักษณะจะคำนวณคะแนนของทั้งคำวิเศษณ์และคำคุณศัพท์

การคำนวณระดับคะแนนคุณลักษณะของประโยคที่อยู่ในรูปเอเอซีที่ใช้ในงานวิจัยนี้ได้ใช้วิธีซึ่งปรับปรุงมาจากวิธีการของเบนมาราและคณะ (Benamara et al., 2006) การคำนวณดังกล่าวจะใช้จากการวัดคะแนนความสัมพันธ์ (r) ระหว่างคำคุณศัพท์และคำวิเศษณ์ในการเชื่อมโยงกันเชิงความหมายในด้านการท่องเที่ยว ซึ่งมีสามรูปแบบ ได้แก่ รูปแบบเอกภาค (Unary Form) รูปแบบทวิภาค (Binary Form) และรูปแบบนิเสธ (Negation Form)

a) รูปแบบเอกภาค จะอยู่ในรูปแบบของ <คำวิเศษณ์> <คำคุณศัพท์>

ตัวอย่างเช่น “really nice” “extremely uncomfortable”

หาก $sc(adj.)$ $sc(adv.)$ และ $sc(AAC)$ แทนคะแนนของคำคุณศัพท์ คำวิเศษณ์ และเอเอซี ตามลำดับ ตัวแปร r คือคะแนนความสัมพันธ์ ซึ่งสามารถคำนวณได้ดังนี้

$$\text{ถ้า } sc(adj.) \geq 3 \text{ แล้ว } sc(AAC) = sc(adj.) + (r * sc(adv.))$$

$$\text{ถ้า } sc(adj.) < 3 \text{ แล้ว } sc(AAC) = sc(adj.) - (r * sc(adv.))$$

ตัวอย่างที่ 1: หากประโยค คือ “The reception staffs are very friendly” กำหนดให้ $r = 0.9$ $sc(\text{friendly}) = 4$ และ $sc(\text{very}) = \pm 1$ ดังนั้นคะแนนคำคุณลักษณะ “reception staffs” สามารถคำนวณได้จากเอเอซี “very friendly” ดังนี้

$$sc(\text{very friendly}) = sc(\text{friendly}) + (r * sc(\text{very}))$$

$$sc(\text{very friendly}) = 4 + (0.9 * 1) = 4.9$$

b) รูปแบบทวีภาค จะอยู่ในรูปแบบของ <คำวิเศษณ์2> <คำวิเศษณ์1> <คำคุณศัพท์>

ตัวอย่างเช่น “really so small” “really very clean”

การคำนวณจะแบ่งเป็นสองขั้นตอน ขั้นตอนแรกคำนวณหา $sc(\text{AAC1})$ จาก Adj. และ Adv.1 หลังจากนั้นนำ $sc(\text{AAC1})$ ไปคำนวณเพื่อหา $sc(\text{AAC2})$ ซึ่งคือคะแนนสุดท้ายของรูปแบบทวีภาค การคำนวณทั้งหมดสามารถสรุปได้ดังนี้

$$\text{ถ้า } sc(\text{adj.}) \geq 3 \text{ แล้ว } sc(\text{AAC1}) = sc(\text{adj.}) + (r_{\text{adv1}} * sc(\text{adv.1}))$$

$$sc(\text{AAC2}) = sc(\text{AAC1}) + (r_{\text{adv2}} * sc(\text{adv.2}))$$

$$\text{ถ้า } sc(\text{adj.}) < 3 \text{ แล้ว } sc(\text{AAC1}) = sc(\text{adj.}) - (r_{\text{adv1}} * sc(\text{adv.1}))$$

$$sc(\text{AAC2}) = sc(\text{AAC1}) - (r_{\text{adv2}} * sc(\text{adv.2}))$$

ตัวอย่างที่ 2: หากประโยค คือ “The reception staffs are really very friendly” กำหนดให้ $sc(\text{really}) = \pm 1$ และคะแนนอื่น ๆ เหมือนกับตัวอย่างที่ 1 ดังนั้นคะแนนคำคุณลักษณะ “reception staff” สามารถคำนวณได้จากเอเอซี “really very friendly” ดังนี้

$$sc(\text{AAC1}) = sc(\text{very friendly}) = sc(\text{friendly}) + (r * sc(\text{very}))$$

$$sc(\text{very friendly}) = 4 + (0.9 * 1) = 4.9$$

$$sc(\text{really very friendly}) = sc(\text{very friendly}) + (r * sc(\text{really}))$$

$$sc(\text{really very friendly}) = 4.9 + (0.9 * 1) = 5.8$$

หมายเหตุ: หากคะแนนคำคุณลักษณะมีค่าเกิน 5 คะแนนดังกล่าวจะถูกปรับเป็น 5

c) รูปแบบนิเสธ จะอยู่ในรูปแบบของ <คำวิเศษณ์นิเสธ> <คำวิเศษณ์> <คำคุณศัพท์>

ตัวอย่างเช่น “not very helpful” “never very clean”

การคำนวณจะแบ่งเป็นสองขั้นตอน ขั้นตอนแรกคำนวณหา $sc(AAC1)$ จาก $nadv.$ และ $adj.$ หลังจากนั้นนำ $sc(AAC1)$ ไปคำนวณเพื่อหา $sc(AAC2)$ ซึ่งเป็นคะแนนสุดท้ายของรูปแบบนิเสธ การคำนวณทั้งหมดสามารถสรุปได้ดังนี้

$$\text{ถ้า } sc(adj.) \geq 3 \text{ แล้ว } sc(AAC1) = sc(adj.) - sc(nadv.)$$

$$\text{ถ้า } sc(AAC1) \geq 3 \text{ แล้ว } sc(AAC2) = sc(AAC1) + (r * sc(adv.))$$

$$\text{ถ้า } sc(AAC1) < 3 \text{ แล้ว } sc(AAC2) = sc(AAC1) - (r * sc(adv.))$$

$$\text{ถ้า } sc(adj.) < 3 \text{ แล้ว } sc(AAC1) = sc(adj.) + sc(nadv.)$$

$$\text{ถ้า } sc(AAC1) \geq 3 \text{ แล้ว } sc(AAC2) = sc(AAC1) + (r * sc(adv.))$$

$$\text{ถ้า } sc(AAC1) < 3 \text{ แล้ว } sc(AAC2) = sc(AAC1) - (r * sc(adv.))$$

ตัวอย่างที่ 3: หากประโยค คือ “The reception staffs are not very friendly” กำหนดให้ $sc(not) = \pm 2$ และคะแนนอื่น ๆ เหมือนกับตัวอย่างที่ 1 ดังนั้นคะแนนคำคุณลักษณะ “reception staff” สามารถคำนวณได้จากเอเอซี “not very friendly” ดังนี้

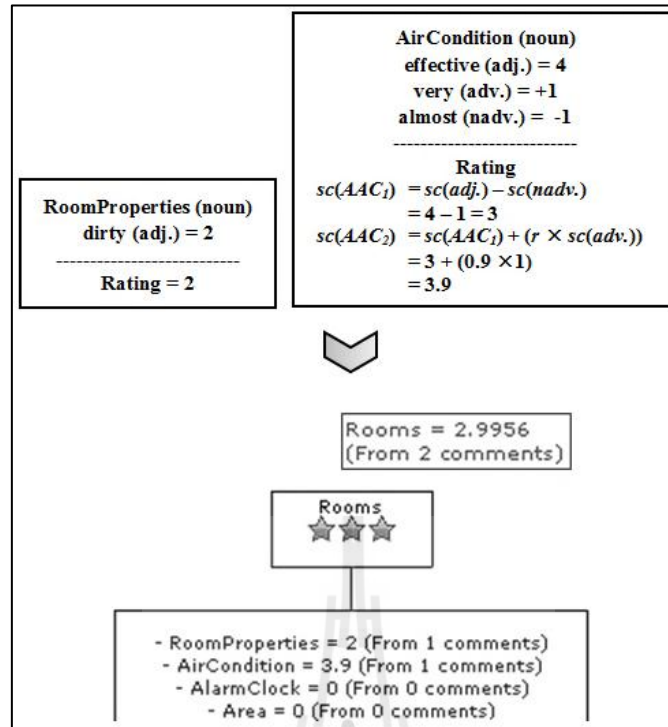
$$sc(AAC1) = sc(not\ friendly) = sc(friendly) - sc(not)$$

$$sc(not\ friendly) = 4 - 2 = 2$$

$$sc(not\ very\ friendly) = sc(not\ friendly) - (r * sc(very))$$

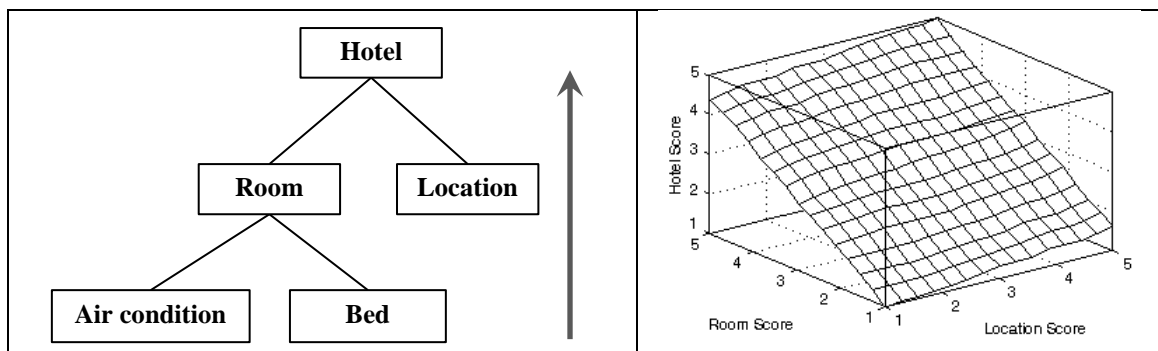
$$sc(not\ very\ friendly) = 2 + (0.9 * 1) = 1.1$$

การคำนวณระดับคะแนนคุณลักษณะของประโยคที่อยู่ในรูปเอเอซีทั้งสามรูปแบบสามารถประยุกต์ใช้กับบทวิจารณ์ที่พักทั้งกรณีประโยคความเดียว (simple sentence) และประโยคความซ้อน (complex sentence) ตัวอย่างของการคำนวณระดับคะแนนคุณลักษณะ “Room” จากประโยคความซ้อน “The room is dirty but the air conditioning was almost very effective” แสดงดังภาพที่ 3.10 แต่ละประโยคจะถูกวิเคราะห์และคำนวณระดับคะแนน เช่น คุณลักษณะ “RoomProperties” จะมีคะแนนเป็น 2 จากประโยคความเดียวประโยคแรก (“The room is dirty”) ในขณะที่คุณลักษณะ “AirCondition” จะมีคะแนนเป็น 3.9 จากประโยคความเดียวประโยคที่สอง (“The air conditioning was almost very effective”)



ภาพที่ 3.10 ตัวอย่างการคำนวณระดับคะแนนคุณลักษณะของที่พัก

2) การคำนวณระดับคะแนนคุณลักษณะของที่พักแบบเป็นลำดับชั้น (Hierarchical Feature Rating) ทำหน้าที่ในการคำนวณระดับคะแนนในภาพรวมของคุณลักษณะของที่พักในลำดับชั้นที่สูงกว่าเรื่อยไปจนกระทั่งถึงลำดับชั้นบนสุด นั่นก็คือ ระดับคะแนนโดยรวมของที่พักแห่งนั้น ตัวอย่างเช่น การคำนวณระดับคะแนนของห้องพัก ได้มาจากการนำระดับคะแนนคุณลักษณะของที่พัก เรื่องเครื่องปรับอากาศ และเรื่องเตียงนอน ที่สกัดมาได้จากกระบวนการก่อนหน้า และหลังจากนั้น ระดับคะแนนของห้องพัก จะถูกนำไปเป็นข้อมูลตั้งต้น ร่วมกับคุณลักษณะเรื่องสถานที่ตั้ง เพื่อคำนวณระดับคะแนนของโรงแรมโดยรวม ซึ่งอยู่ในลำดับชั้นบนสุด ดังแสดงในภาพที่ 3.11 ทั้งนี้ ความรู้ที่สกัดได้ทั้งหมดจะถูกจัดเก็บลงในฐานความรู้ของระบบเพื่อใช้ในการสืบค้นต่อไป



ภาพที่ 3.11 ตัวอย่างการคำนวณระดับคะแนนคุณลักษณะของที่พักแบบเป็นลำดับชั้น

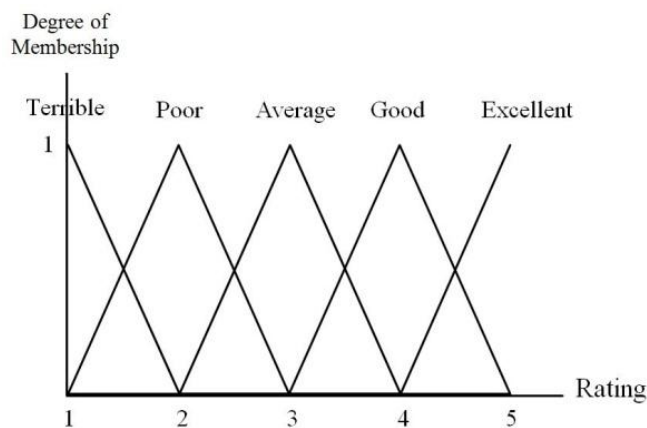
ทั้งนี้ ระบบวินิจฉัยแบบฟัซซีที่นำมาใช้ในการคำนวณระดับคะแนนในภาพรวมนั้น มีส่วนประกอบพื้นฐานที่สำคัญอยู่ 4 ส่วน ดังนี้

- ส่วนฟัซซีฟิเคชัน (Fuzzification)

ฟัซซีฟิเคชันเป็นการนำเอาค่าของระดับคะแนนเฉพาะเรื่องมาแปลงให้เป็นค่าแบบฟัซซี (Fuzzy Values) ซึ่งกระบวนการนี้ต้องมีการกำหนดฟัซซีเซตสำหรับคุณสมบัติของโรงเรียนที่มีอยู่ในออนไลน์ที่ทั้ง 76 คุณสมบัติ ซึ่งฟัซซีเซตของแต่ละคุณสมบัติได้กำหนดรูปแบบของระดับค่าความเป็นสมาชิก (Degree of membership) แบบฟังก์ชันสามเหลี่ยม (Triangular Membership Function) ซึ่งประกอบด้วยพารามิเตอร์ 3 ค่า ได้แก่ {a, b, c} ดังสมการที่ 3-1 (Nguyen and Walker, 2000)

$$Triangular(x: a, b, c) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \leq x \leq b \\ (c-x)/(c-b) & b \leq x \leq c \\ 0 & x > c \end{cases} \quad (3-1)$$

ฟังก์ชันสามเหลี่ยมนี้ได้ถูกนำมาใช้แบ่งกลุ่มสมาชิกของตัวแปรระดับคะแนน (Rating) ออกเป็น 5 กลุ่มสมาชิก ได้แก่ แย่มาก (Terrible) แย่ (Poor) ปานกลาง (Average) ดี (Good) และดีมาก (Excellent) โดยกำหนดค่าพารามิเตอร์ในแต่ละฟังก์ชันสมาชิกแตกต่างกัน อาทิ การแบ่งกลุ่มสมาชิกตัวแปรที่มีคะแนนอยู่ในระดับแย่ (Poor) กำหนดค่าพารามิเตอร์ให้ a=1 b=2 และ c=3 เป็นต้น ทั้งนี้ การแบ่งกลุ่มสมาชิกของตัวแปรระดับคะแนน แสดงดังภาพที่ 3.12



ภาพที่ 3.12 การแบ่งกลุ่มสมาชิกตัวแปรระดับคะแนน

- ส่วนฐานกฎฟuzzy (Fuzzy Rule Base)

ฐานกฎฟuzzy ประกอบด้วยกฎฟuzzyต่าง ๆ ที่ใช้ในการอนุมานจากเหตุไปสู่ผลลัพธ์ นั่นก็คือ หากนักท่องเที่ยววิจารณ์เกี่ยวกับคุณสมบัติต่าง ๆ ของโรงแรม จะอนุมานหรือสรุประดับคะแนนในภาพรวมของโรงแรมได้ ซึ่งนิยม

เขียนแสดงความสัมพันธ์อยู่ในรูปของ ถ้า-แล้ว (If-Then) อาทิ “If Location=Good Then Hotel=Good”

ในการคำนวณระดับคะแนนในภาพรวมโดยใช้ตรรกศาสตร์คลุมเครือนี้ ผู้วิจัยได้สร้างกฎขึ้นมาจำนวนทั้งสิ้น 395 กฎ แบ่งเป็นกฎของฟuzzyเซตเรื่องห้องน้ำ 40 กฎ ห้องพัก 135 กฎ การบริการ 25 กฎ และภาพรวมของทั้งโรงแรม 195 กฎ ซึ่งกฎของฟuzzyเซตแต่ละเรื่องนี้ จะมีการให้ค่าน้ำหนักของแต่ละกฎที่แตกต่างกันไป โดยการกำหนดน้ำหนักของแต่ละกฎจะเป็นไปตามแนวคิดที่ว่า **“คุณสมบัติของโรงแรมที่กำหนดในแต่ละกฎนั้น สอดคล้องกับระดับคะแนนในภาพรวมของทั้งโรงแรมมากน้อยเพียงไร”** ตัวอย่างเช่น เมื่อนักท่องเที่ยววิจารณ์ว่าโรงแรมนี้มีทำเลที่ตั้งดี (คิดเป็นระดับ 4 คะแนน) แต่คะแนนในภาพรวมของบทวิจารณ์ถูกกำหนดว่าโรงแรมนี้ได้เพียงระดับที่ 1 คือแย่มาก แสดงว่าคุณสมบัติเรื่องทำเลที่ตั้งสอดคล้องกับระดับคะแนนในภาพรวมของทั้งโรงแรมน้อยมาก ค่าน้ำหนักของกฎเรื่องทำเลที่ตั้งจึงควรกำหนดให้เหมาะสม ในที่นี้ค่าน้ำหนักของกฎจะกำหนดให้มีค่า 3 ค่า คือ 0.1 0.5 และ 1 ซึ่งค่าน้ำหนักมาก (เท่ากับ 1) หมายถึง การวิจารณ์ในเรื่องนั้นสอดคล้องกับระดับคะแนนในภาพรวมมาก ค่าน้ำหนักปานกลาง (เท่ากับ 0.5) หมายถึง การวิจารณ์ในเรื่องนั้นสอดคล้องกับระดับคะแนนในภาพรวมปานกลาง และค่าน้ำหนักน้อย (เท่ากับ 0.1) หมายถึง การวิจารณ์ในเรื่องนั้นสอดคล้องกับระดับคะแนนในภาพรวมน้อย

โดยการพิจารณาว่าคุณสมบัติต่าง ๆ สอดคล้องกับระดับคะแนนในภาพรวมของทั้งโรงแรมมากน้อยเพียงไร จะได้มาจากการวิเคราะห์การให้คะแนนของกลุ่มตัวอย่าง ซึ่งก็คือบทวิจารณ์ที่สุ่มเก็บมาจากเว็บไซต์ TripAdvisor.com จำนวน 100 บทวิจารณ์ โดยมีระดับความพึงพอใจในภาพรวมของทั้งโรงแรมตามที่นักท่องเที่ยวให้คะแนนแตกต่างกัน 5 ระดับคะแนนจำนวนระดับคะแนนละ 20 บทวิจารณ์ โดยพิจารณาว่า คะแนนเฉลี่ยในเรื่องต่าง ๆ ที่นักท่องเที่ยวกำหนดให้ในแต่ละระดับ มีความใกล้เคียงกับคะแนนในภาพรวมของทั้งโรงแรมที่ระดับ

ในตารางที่ 3.2 เป็นข้อมูลการให้คะแนนของนักท่องเที่ยวจากกลุ่มตัวอย่างบทวิจารณ์ ในเรื่องทำเลที่ตั้ง พบว่า มี 3 บทวิจารณ์ที่กล่าวถึงทำเลที่ตั้งของโรงแรมระดับที่ 1 โดยนักท่องเที่ยวให้คะแนนในภาพรวมของทั้งโรงแรมแย่มาก (ระดับที่ 1) แต่วิจารณ์ว่าโรงแรมนั้นมีทำเลที่ตั้งดี (โดยเฉลี่ย 3 บทวิจารณ์ได้ระดับคะแนนเฉพาะเรื่องทำเลที่ตั้ง 4 คะแนน) เป็นต้น

ตารางที่ 3.2 ตัวอย่างการให้คะแนนของนักท่องเที่ยวในเรื่องทำเลที่ตั้งจากกลุ่มตัวอย่างบทวิจารณ์

ระดับคะแนน ในภาพรวมของทั้งโรงแรม	ระดับคะแนน เรื่องทำเลที่ตั้ง
ระดับที่ 1 ★	พบการวิจารณ์ 3 ครั้ง (คะแนนเฉลี่ย 4 คะแนน)
ระดับที่ 2 ★★	พบการวิจารณ์ 6 ครั้ง (คะแนนเฉลี่ย 4 คะแนน)
ระดับที่ 3 ★★★	พบการวิจารณ์ 10 ครั้ง (คะแนนเฉลี่ย 4.4 คะแนน)
ระดับที่ 4 ★★★★	พบการวิจารณ์ 13 ครั้ง (คะแนนเฉลี่ย 4.31 คะแนน)
ระดับที่ 5 ★★★★★	พบการวิจารณ์ 7 ครั้ง (คะแนนเฉลี่ย 4.57 คะแนน)

จากตารางที่ 3.2 แสดงให้เห็นว่า ผู้ใช้ให้คะแนนในเรื่องทำเลที่ตั้งมีความใกล้เคียงกับคะแนนในภาพรวมของทั้งโรงแรมจำนวน 2 ระดับ (คือ ระดับ 4 และระดับ 5) แต่ในระดับ 1 2 และ 3 มีความแตกต่างกับการวิจารณ์เรื่องทำเลที่ตั้ง ตามเกณฑ์ความใกล้เคียงที่กำหนดไว้ในตารางที่ 3.3

ตารางที่ 3.3 เกณฑ์การพิจารณาความใกล้เคียงระหว่างระดับคะแนนในเรื่องต่าง ๆ กับระดับคะแนนในภาพรวมของทั้งโรงแรม

เกณฑ์ความใกล้เคียง	
คะแนนเฉลี่ย 1.00-1.99	กำหนดว่าให้ใกล้เคียงกับระดับที่ 1
คะแนนเฉลี่ย 1.01-2.99	กำหนดว่าให้ใกล้เคียงกับระดับที่ 2
คะแนนเฉลี่ย 2.01-3.99	กำหนดว่าให้ใกล้เคียงกับระดับที่ 3
คะแนนเฉลี่ย 3.01-4.99	กำหนดว่าให้ใกล้เคียงกับระดับที่ 4
คะแนนเฉลี่ย 4.01-5.00	กำหนดว่าให้ใกล้เคียงกับระดับที่ 5

จากการพิจารณาว่า คะแนนเฉลี่ยในเรื่องต่าง ๆ มีความใกล้เคียงกับคะแนนในภาพรวมของทั้งโรงแรมกี่ระดับ จะนำมาคำนวณเป็นร้อยละความน่าจะเป็นของการวิจารณ์เรื่องนั้น ๆ ว่าสอดคล้องกับ

ระดับคะแนนในภาพรวมของทั้งโรงแรมเท่าใด ซึ่งจากการวิเคราะห์เรื่องทำเลที่ตั้งจะดูว่ามีค่าความน่าจะเป็นเพียง 40% ($2/5 \times 100$) เนื่องจากมีความใกล้เคียงเพียง 2 ระดับจาก 5 ระดับ

จากนั้นจึงนำค่าร้อยละความน่าจะเป็นดังกล่าว มาแปลงเป็นค่าน้ำหนักของกฎตามเกณฑ์ที่แสดงในตารางที่ 3.4

ตารางที่ 3.4 เกณฑ์การพิจารณาค่าน้ำหนักจากร้อยละความน่าจะเป็นที่การวิจารณ์เฉพาะเรื่องหนึ่ง ๆ จะส่งผลต่อระดับคะแนนในภาพรวม

ร้อยละความน่าจะเป็น	ค่าน้ำหนักของกฎ
มากกว่าหรือเท่ากับ 80%	1
มากกว่าหรือเท่ากับ 50%	0.5
น้อยกว่า 50%	0.1

ในตารางที่ 3.5 เป็นตัวอย่างของกฎตรรกศาสตร์คลุมเครือ 13 กฎจากทั้งหมด 395 กฎ ซึ่งสังเกตเห็นว่าแต่ละกฎมีตัวแปรนำเข้าเพียงตัวแปรเดียว เนื่องจากการวิจัยนี้ได้สร้างกฎโดยแยกพิจารณาการสอดคล้องกับระดับคะแนนในภาพรวมของทั้งโรงแรมเป็นเรื่อง ๆ ทำให้แต่ละเรื่องมีค่าน้ำหนักของกฎเฉพาะของตนเอง และอาจไม่เท่ากับเรื่องอื่น ๆ ดังนั้น จึงได้สร้างกฎแต่ละกฎในลักษณะที่มีตัวแปรนำเข้าเพียงตัวแปรเดียว เพื่อให้กฎที่ได้มีความเหมาะสมต่อการใช้งานด้านการวิเคราะห์บทวิจารณ์มากที่สุด

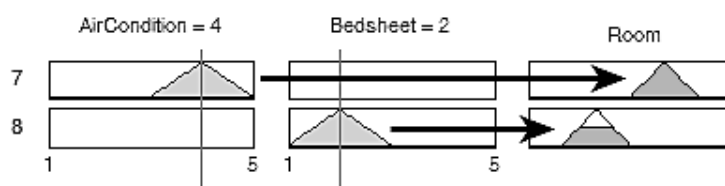
- ส่วนกลไกการอนุมาน (Inference Engine)

กลไกการอนุมานเป็นส่วนที่ใช้วินิจฉัยเพื่อตัดสินใจระดับคะแนนในภาพรวมตามกฎที่มีอยู่แบบลำดับชั้น ซึ่งจากกฎและน้ำหนักที่ได้ในฐานกฎฟัซซีจะนำมาใช้เป็นเกณฑ์การคำนวณระดับคะแนนในภาพรวม โดยพิจารณาระดับคะแนนของทุกเรื่องในขั้นก่อนหน้าเป็นข้อมูลนำเข้า และระดับคะแนนในภาพรวมเป็นข้อมูลผลลัพธ์ โดยในลำดับชั้นแรกกลไกการอนุมานจะพิจารณาระดับคะแนนในภาพรวมจากกฎของฟัซซีเซตในกลุ่มเรื่องห้องน้ำ ห้องพัก และการบริการก่อน จากนั้นจึงนำระดับคะแนนในภาพรวมที่ได้จากลำดับชั้นแรกนี้มาหาค่าคะแนนในภาพรวมของทั้งโรงแรมต่อไป

ตารางที่ 3.5 ตัวอย่างกฎตรรกศาสตร์คลุมเครือที่ใช้ในการคำนวณระดับคะแนนในภาพรวม

Rules	If	Then	Weight
- กฎที่ใช้คำนวณคะแนนในภาพรวมของโรงแรม			
1	Room = Average	Hotel = Average	1
2	Room = Good	Hotel = Good	1
3	Service = Poor	Hotel = Poor	1
4	Bathroom = Average	Hotel = Average	1
5	Value = Excellent	Hotel = Excellent	0.5
6	Location = Excellent	Hotel = Excellent	0.1
- กฎที่ใช้คำนวณคะแนนในรายละเอียดของโรงแรม			
7	AirCondition = Good	Room = Good	1
8	Bedsheet = Bad	Room = Bad	0.5
9	TVSet = Average	Room = Average	0.1
10	Employee = Poor	Service = Poor	0.5
11	Reception = Excellent	Service = Excellent	0.1
12	Bathtub = Terrible	Bathroom = Terrible	0.5
13	Toilet = Good	Bathroom = Good	0.1

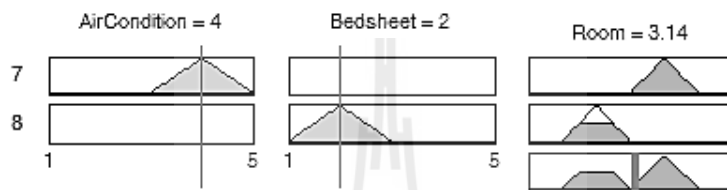
ในโลกการอนุมาน หากคะแนนที่เป็นข้อมูลนำเข้ามีการเปลี่ยนแปลง คะแนนของข้อมูลผลลัพธ์ก็จะเปลี่ยนตามไปด้วย ทั้งนี้ การเปลี่ยนแปลงจะเกิดขึ้นมากหรือน้อยขึ้นอยู่กับน้ำหนักของกฎที่เกี่ยวข้องกับข้อมูลผลลัพธ์นั้น ตัวอย่างเช่น การคำนวณหาระดับคะแนนในภาพรวมของห้องพักที่พิจารณาจากข้อมูลนำเข้า คือคะแนนของเครื่องปรับอากาศ (AirCondition) จำนวน 4 คะแนน และผ้าปูเตียง (Bedsheet) จำนวน 2 คะแนน เมื่อนำมาอนุมานจะไปเรียกใช้กฎที่ 7 และกฎที่ 8 ดังแสดงในรูปที่ 6 จะเห็นได้ว่ากฎที่ 7 “If AirCondition = Good Then Room = Good” มีค่าน้ำหนักเท่ากับ 1 ทำให้ระดับความเป็นสมาชิกของ “Room=Good” มีค่าสูงสุด ดังแสดงในภาพด้านขวาของภาพที่ 3.13 ในขณะที่กฎที่ 8 “If Bedsheet = Bad Then Room = Bad” มีค่าน้ำหนักเท่ากับ 0.5 ทำให้ระดับความเป็นสมาชิกของ “Room=Bad” มีค่าเพียงครึ่งหนึ่ง เป็นต้น



ภาพที่ 3.13 ตัวอย่างการอนุมานค่าผลลัพธ์ เมื่อค่าน้ำหนักของกฎเป็น 1 และ 0.5

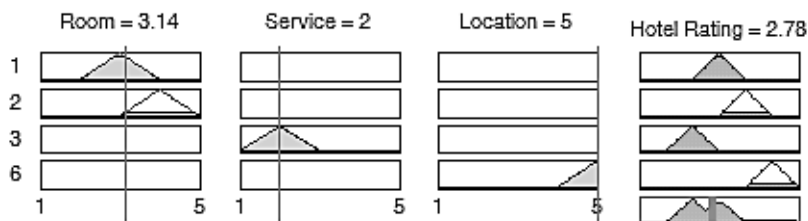
- ส่วนดีฟัซซิฟิเคชัน (Defuzzification)

ดีฟัซซิฟิเคชันเป็นส่วนที่รวมผลลัพธ์ที่ได้จากการทำงานของกฎต่าง ๆ และแปลงให้เป็นระดับคะแนนในภาพรวม โดยในงานวิจัยนี้เลือกใช้วิธีเซนทรอยด์ (Centroid of Area Method: COA) (Jang, Sun and Mizutani, 1997) ในการคำนวณระดับคะแนนในภาพรวม ยกตัวอย่างการคำนวณระดับคะแนนในภาพรวมของเรื่องห้องพัก จากการพิจารณาจากคะแนนของเครื่องปรับอากาศจำนวน 4 คะแนน และผ้าปูเตียงจำนวน 2 คะแนน จะได้ผลลัพธ์คือ ห้องพักมีระดับคะแนนเท่ากับ 3.14 คะแนน จากการรวมผลลัพธ์ของกฎที่ 7 และกฎที่ 8 ดังแสดงในภาพที่ 3.14



ภาพที่ 3.14 ตัวอย่างการคำนวณระดับคะแนนในภาพรวมของเรื่องห้องพัก เมื่อคะแนนเรื่องเครื่องปรับอากาศเป็น 4 คะแนน และเรื่องผ้าปูเตียงเป็น 2 คะแนน

ทั้งนี้ เมื่อนำระดับคะแนนในเรื่องห้องพักจำนวน 3.14 คะแนน มาคำนวณหาระดับคะแนนในภาพรวมของทั้งโรงแรม โดยพิจารณาจากคะแนนของการบริการจำนวน 2 คะแนน และเรื่องทำเลที่ตั้งจำนวน 5 คะแนนร่วมด้วย จะได้ผลลัพธ์เป็นคะแนนในภาพรวมของทั้งโรงแรมมีค่า 2.78 คะแนน ดังแสดงในภาพที่ 3.15 ทั้งนี้จะสังเกตเห็นว่าระดับคะแนนในเรื่องทำเลที่ตั้งสอดคล้องกับคะแนนในภาพรวมของทั้งโรงแรมน้อยมาก เนื่องจากค่าน้ำหนักของกฎในเรื่องนี้มีค่าเพียง 0.1 ตามกฎที่ 6 ในตารางที่ 3.5 ซึ่งแตกต่างจากกฎที่ 1-3 ที่มีค่าน้ำหนักของกฎสูงสุดคือ 1



ภาพที่ 3.15 ตัวอย่างการคำนวณระดับคะแนนในภาพรวมของทั้งโรงแรม เมื่อคะแนนเรื่องห้องพักเป็น 3.14 คะแนน เรื่องการบริการเป็น 2 คะแนน และเรื่องทำเลที่ตั้งเป็น 5 คะแนน

3.1.3.4 ส่วนอธิบายความรู้

สำหรับในบริบทของงานวิจัยนี้ ความรู้ หมายถึง ข้อมูลสารสนเทศที่ผสมผสานกับประสบการณ์ ความรอบรู้ในบริบท การแปลความหมายและการแสดงความคิดเห็น โดยเป็นข้อมูลสารสนเทศที่มีคุณค่าสูงที่พร้อมจะนำไปประยุกต์ใช้ในการตัดสินใจ หรือใช้ในการทำงาน (Davenport and Prusak, 1998) โดยรูปแบบความรู้ของระบบจะแสดงตามเรื่องที่นักท่องเที่ยวได้แสดงความคิดเห็นหรือมีความสนใจ เช่น การบริการ บรรยากาศสภาพแวดล้อม ห้องพัก อาหาร กิจกรรม ราคา ความปลอดภัย ความสะอาด ความสะดวกในเดินทาง (Hsu, Tsai and Wu, 2009: 295; Cracolici and Nijkamp, 2008: 342; Chi and Qu, 2008: 629-631; Del Bosque and San Martin, 2008: 562) พร้อมทั้งแสดงระดับคะแนนในเรื่องนั้น ๆ ซึ่งได้จากความรู้ที่สกัดได้มาจัดให้อยู่ในรูปแบบที่ผู้ใช้สามารถเข้าใจได้ง่ายและชัดเจนมากยิ่งขึ้น โดยความรู้ของระบบจะแสดงเป็นกราฟต้นไม้ของที่พักแห่งหนึ่งตามรายละเอียดของสถานที่พักที่นักท่องเที่ยวได้แสดงความคิดเห็นไว้ และแบ่งตามหมวดหมู่ที่ได้กำหนด เช่น ทำเลที่ตั้ง การบริการ รายละเอียดของการบริการ ห้องพัก และรายละเอียดของที่พัก เป็นต้น นอกจากนี้ กราฟต้นไม้ยังแสดงระดับคะแนนและจำนวนคำวิจารณ์ของเรื่องต่าง ๆ ที่สกัดได้อย่างเป็นลำดับชั้น โดยระบบจะแสดงผลระดับคะแนนโดยรวมเป็นรูปดาวในจำนวนที่แตกต่างกันตามเศษทศนิยม ซึ่งมีเกณฑ์ดังต่อไปนี้

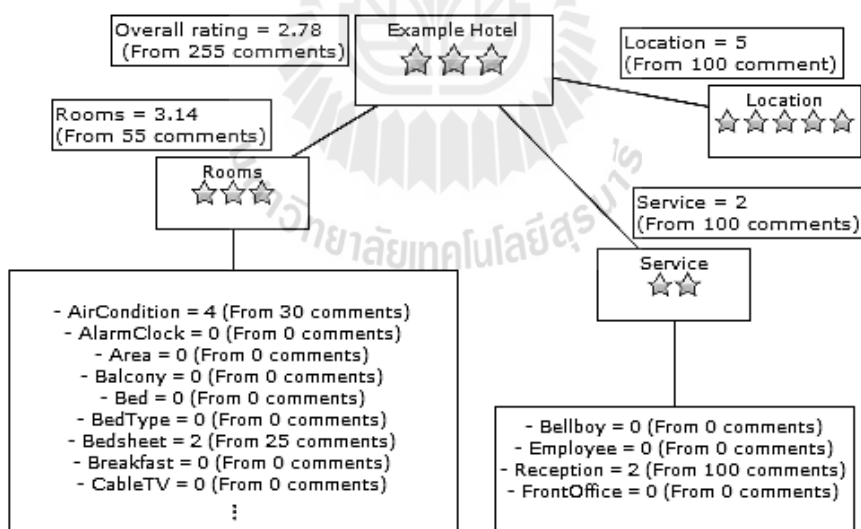
- ถ้าเศษทศนิยมมีค่าอยู่ในช่วง 0.00-0.24
จะแสดงรูปดาวจำนวนเท่ากับระดับคะแนนที่คำนวณได้
- ถ้าเศษทศนิยมมีค่าอยู่ในช่วง 0.25-0.74
จะแสดงรูปดาวจำนวนเท่ากับระดับคะแนนที่คำนวณได้ และเพิ่มรูปดาวอีกครั้งดวง
- ถ้าเศษทศนิยมมีค่าอยู่ในช่วง 0.75-0.99
จะแสดงรูปดาวจำนวนเพิ่มขึ้น 1 ดวงจากระดับคะแนนที่คำนวณได้

ตัวอย่างการแสดงผลระดับคะแนนโดยรวมเป็นรูปดาว เมื่อจำนวนเต็มของระดับคะแนนทั้งที่พักเท่ากับ 4 คะแนนแสดงดังตารางที่ 3.6

ตารางที่ 3.6 ตัวอย่างการแสดงผลระดับคะแนนเป็นรูปดาว เมื่อจำนวนเต็มของระดับคะแนนทั้งโรงแรม เท่ากับ 4 คะแนน

เศษทศนิยม	จำนวนรูปดาวที่แสดง
4.00-4.24	★★★★
4.25-4.74	★★★★½
4.75-4.99	★★★★★

จากภาพที่ 3.16 ได้แสดงรายละเอียดของโรงแรมแห่งหนึ่งว่าเป็นโรงแรม 3 ดาว (ระดับคะแนน 2.78) และมีการวิจารณ์ของผู้ที่เคยไปมาแล้วทั้งหมด 255 คำวิจารณ์ โดยระดับคะแนน 2.78 นั้นเกิดจากการแสดงความคิดเห็นในด้านต่าง ๆ ได้แก่ 1) ทำเลที่ตั้ง 5 คะแนน จาก 100 คำวิจารณ์; 2) การบริการ 2 คะแนน จาก 100 คำวิจารณ์; และ 3) ห้องพัก 3.14 คะแนน จาก 55 คำวิจารณ์ ซึ่งการอธิบายด้วยกราฟเพื่อแสดงรายละเอียดของโรงแรมนี้จะเป็นประโยชน์อย่างยิ่งต่อการตัดสินใจเลือกโรงแรมให้ตรงกับความต้องการของนักท่องเที่ยวได้มากที่สุด



1=Terrible 2=Poor 3=Average 4=Good 5=Excellent

ภาพที่ 3.16 ตัวอย่างการแสดงผลความรู้ที่สกัดได้

จากการออกแบบและพัฒนาสามารถสรุปได้ว่า งานวิจัยนี้ได้พัฒนาระบบสกัดและค้นคืนความรู้แบบอัตโนมัติจากบทวิจารณ์ออนไลน์ของนักท่องเที่ยว โดยนำเอาเทคนิคต่าง ๆ ที่มีอยู่

ในปัจจุบัน เช่น ตัวแปลภาษา ออนโทโลยี และตรรกศาสตร์คลุมเครือ มาผสมผสานเพื่อให้สามารถสกัดและค้นคืนความรู้ได้อย่างมีประสิทธิภาพมากที่สุด

3.1.4 ทดสอบและประเมินผลการทำงานของระบบ

การทดสอบการทำงานของระบบและประเมินผลจะทำโดยการนำบทวิจารณ์ของนักท่งเกี่ยวข้องมาเป็นข้อมูลนำเข้า เพื่อตรวจสอบการทำงานของระบบ ตลอดจนความรู้ที่ระบบสกัดได้ให้เป็นไปอย่างถูกต้อง โดยในการประเมินผลจะพิจารณาจากความสามารถของระบบว่า สามารถสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์ที่รับเข้ามาได้ถูกต้องหรือไม่ และสามารถคำนวณระดับคะแนนของที่พักได้ใกล้เคียงกับที่นักท่งเกี่ยวข้องเคยให้ไว้หรือไม่ นอกจากนี้ยังมีการทดสอบเพื่อหาค่าระดับคะแนนความสัมพันธ์ระหว่างคำวิเศษณ์และคำคุณศัพท์ (r) ที่เหมาะสมกับด้านการท่งเกี่ยวข้อง โดยการเปรียบเทียบค่าระดับคะแนนที่ระบบคำนวณได้กับค่าระดับคะแนนที่นักท่งเกี่ยวข้องกำหนด

โดยการวัดความถูกต้องของกระบวนการสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์นั้น จะใช้วิธีวัดค่าที่เป็นมาตรฐานในการประเมิน ซึ่งในที่นี้มี 3 ค่า ได้แก่ ค่าความถูกต้อง (*Accuracy*) ค่าความแม่นยำ (*Precision*) และค่าความระลึก (*Recall*) (Miao, Duan, Zhang and Jiao, 2009: 9172)

ส่วนการทดสอบกระบวนการคำนวณระดับคะแนนของที่พัก และการทดสอบเพื่อหาค่าระดับคะแนนความสัมพันธ์ระหว่างคำวิเศษณ์และคำคุณศัพท์ (r) ได้ใช้วิธีวัดค่ามาตรฐานในการประเมินความผิดพลาด 2 ค่า ได้แก่ ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (*MAE*) และค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (*RMSE*) (Memmedli and Ozdemi, 2010: 12-21) ซึ่งจะกล่าวอย่างละเอียดในหัวข้อการวิเคราะห์ข้อมูล

3.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยเพื่อทดสอบความถูกต้องในการสกัดความรู้ มีรายละเอียดดังนี้

3.2.1 เครื่องมือที่ใช้ในการพัฒนาระบบ

1) เครื่องคอมพิวเตอร์เพื่อใช้ในการพัฒนาระบบสกัดความรู้แบบอัตโนมัติจากบทวิจารณ์ออนไลน์ของลูกค้า และการจัดการฐานความรู้ มีคุณสมบัติดังนี้

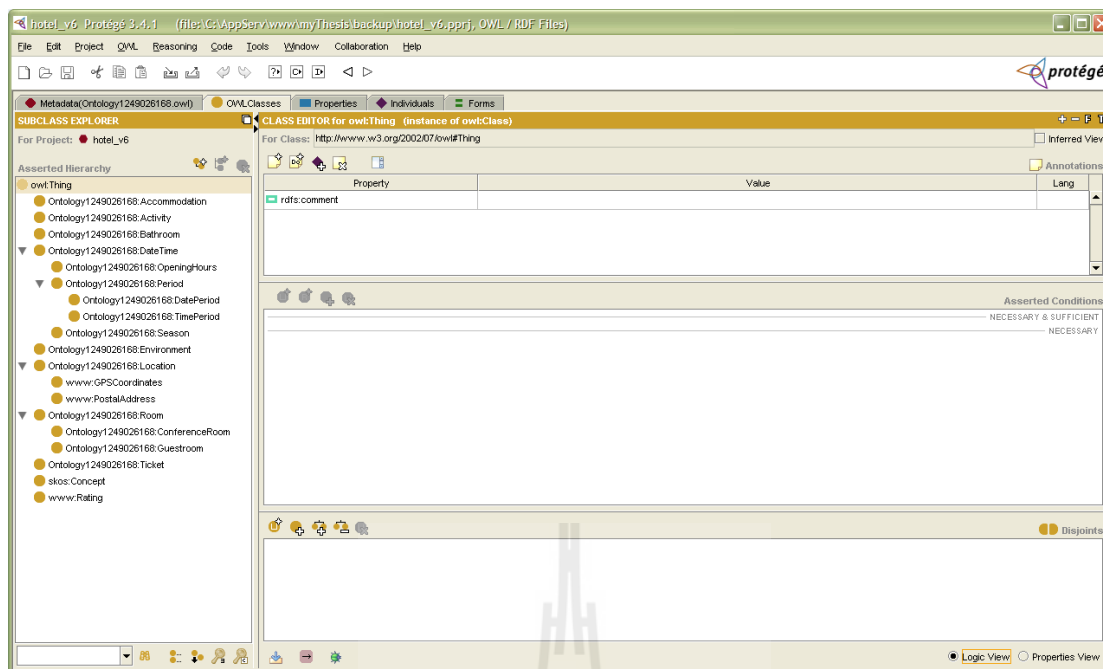
- หน่วยประมวลผลกลางชนิด Intel Centrino Duo Processor ความถี่ 1.6 GHz
- หน่วยความจำหลัก ขนาด 80 GB
- หน่วยความจำสำรอง ขนาด 2.5 GB
- สามารถติดต่อกับระบบเครือข่ายอินเทอร์เน็ตได้
- อุปกรณ์ต่อพ่วงอื่น เช่น เมาส์ แป้นพิมพ์ เครื่องพิมพ์

2) ระบบปฏิบัติการและโปรแกรมประยุกต์ที่ใช้สำหรับการพัฒนาเว็บไซต์ของระบบ โดยมีความโดดเด่นด้านการสร้างโปรแกรมประยุกต์บนอินเทอร์เน็ต และสามารถเชื่อมต่อกับฐานข้อมูลได้

- ระบบปฏิบัติการ: Microsoft Windows XP Version 2002 Service Pack 3
- เว็บเซิร์ฟเวอร์: Apache Web Server 2.2.8
- เว็บเบราว์เซอร์: Windows Internet Explorer 8
- เครื่องมือในการพัฒนา: PHP Script Language 5.2.6
- ฐานข้อมูล: MySQL 5.0.51b
- ระบบจัดการฐานข้อมูล: phpMyAdmin 2.10.3
- โปรแกรมพัฒนาและปรับปรุงออนโทโลยี: Protégé 3.4.1
- โปรแกรมเชื่อมต่อระหว่างระบบกับออนโทโลยีที่พัฒนาขึ้น: RAP (RDF API for PHP)
- โปรแกรมแสดงผลแผนภาพโครงสร้างต้นไม้: PowerCharts Version 3

3.2.2 เครื่องมือในการพัฒนาออนโทโลยี

การสร้างออนโทโลยีซึ่งเป็นฐานความรู้ของระบบนั้น ผู้วิจัยใช้โปรแกรม Protégé รุ่น 3.4.1 ซึ่งเป็นซอฟต์แวร์ในการพัฒนาออนโทโลยี ดังแสดงรูปแบบหน้าจอในภาพที่ 3.17 พัฒนาขึ้นโดยมหาวิทยาลัยสแตนฟอร์ดที่มีผู้ใช้ทั่วโลกทดลองใช้งาน โปรแกรมนี้เตรียมพร้อมเรื่องกราฟิก การปฏิสัมพันธ์การออกแบบออนโทโลยี และสภาพแวดล้อมการพัฒนาฐานความรู้ ผู้ใช้สามารถเข้าถึงประเด็นข้อมูลและจัดการออนโทโลยีได้อย่างรวดเร็วตามความต้องการ ข้อดีข้อหนึ่งของโปรแกรม Protégé คือเป็นระบบเปิด ช่วยให้สร้างระบบโดยเพิ่มฟังก์ชันการทำงานใหม่โดยการสร้างโปรแกรมเสริม (Plug-in) ที่เหมาะสม ซึ่งมีการพัฒนาร่วมกันจากนักพัฒนาทั่วโลก ในการปรับปรุงและพัฒนาออนโทโลยีเพื่อเป็นฐานความรู้ของระบบ โดยครอบคลุมการทำงานในเรื่องการสร้างและบันทึกโครงการงาน (Project) การสร้างคลาส (Classes) การสร้างสล็อต (Slots) และการสร้างอินสแตนซ์ (Instances) ซึ่งมีหน้าจอและลักษณะการใช้งานโปรแกรมดังอธิบายในภาคผนวก ข



ภาพที่ 3.17 หน้าจอหลักของโปรแกรม Protégé

3.3 การเก็บรวบรวมข้อมูล

ข้อมูลบทวิจารณ์ที่นำมาทดสอบจะเป็นคนละชุดกับข้อมูลที่ใช้ในการสร้างขั้นตอนวิธีการสกัดความรู้ โดยผู้วิจัยได้สุ่มเก็บบทวิจารณ์จำนวน 100 บทวิจารณ์ จากที่พักจำนวน 5 ที่ ซึ่งอยู่ในระดับความพึงพอใจที่ต่างกัน 5 ระดับ จำนวนระดับละ 20 บทวิจารณ์ จากเว็บไซต์เพื่อการท่องเที่ยว TripAdvisor.com โดยบทวิจารณ์เหล่านี้จำแนกได้เป็นข้อมูลคำวิจารณ์ที่เกี่ยวข้องกับเรื่องที่พักจำนวน 205 คำวิจารณ์ ข้อมูลคำวิจารณ์ที่ไม่เกี่ยวข้องกับเรื่องที่พักจำนวน 313 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ผิดไวยากรณ์จำนวน 143 คำวิจารณ์

3.4 การวิเคราะห์ข้อมูล

ดังที่กล่าวมาแล้วว่า ส่วนอนุมานความรู้เป็นส่วนสำคัญที่สุดของระบบ ซึ่งประกอบด้วย 2 กระบวนการหลัก ได้แก่ กระบวนการสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์ (Feature Extraction) และกระบวนการคำนวณระดับคะแนนของที่พัก (Accommodation Rating) ดังนั้นในงานวิจัยนี้จึงได้ประเมินความถูกต้องของกระบวนการเหล่านี้ ซึ่งการวิเคราะห์ผลลัพธ์ที่ได้จากการประเมินจะแบ่งตามกระบวนการดังนี้

3.4.1 การวิเคราะห์ผลการประเมินกระบวนการสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์

โดยทั่วไป งานวิจัยทางด้านการสกัดความรู้จะจำแนกข้อมูลที่ใช้ในการประเมินแบ่งออกเป็น 4 กลุ่ม อันได้แก่ ข้อมูลที่สามารถสกัดได้และมีความเกี่ยวข้อง (True Positive: TP), ข้อมูลที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (False Positive: FP), ข้อมูลที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (True Negative: TN) และข้อมูลที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (False Negative: FN) (Miao et al., 2009: 9172) ดังแสดงในตารางที่ 3.7

ตารางที่ 3.7 ข้อมูล 4 ประเภทที่แตกต่างกันตามความเกี่ยวข้องและการสกัดได้

	ข้อมูลที่มีความเกี่ยวข้อง (+)	ข้อมูลที่ไม่มีความเกี่ยวข้อง (-)
ข้อมูลที่สกัดได้ (+)	TP	FP
ข้อมูลที่สกัดไม่ได้ (-)	FN	TN

ในการประเมินความถูกต้องของการสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์นั้น ข้อมูลที่นำมาทดสอบจะเป็นคนละชุดกับข้อมูลที่นำมาใช้ในการสร้างขั้นตอนวิธีในการสกัดความรู้ โดยในการประเมินความถูกต้องนี้จะทำโดยการตรวจสอบความรู้ที่สกัดได้จากระบบในแง่ความใกล้เคียงของระดับคะแนนที่ระบบสกัดได้กับระดับคะแนนโดยรวมที่ผู้กำหนด และนำผลการวิเคราะห์ประโยคคำวิจารณ์ที่ได้มาเป็นเกณฑ์ในการวัดความถูกต้องเทียบกับขั้นตอนวิธีในการสกัดความรู้แบบอัตโนมัติที่ได้พัฒนาขึ้นมา โดยการทดสอบความถูกต้องนั้นจะใช้วิธีวัดค่าที่เป็นมาตรฐานในการประเมิน ซึ่งในที่นี้จะมี 3 ค่า ได้แก่ ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) โดยมีสูตรดังสมการที่ 3-2 3-3 และ 3-4 ตามลำดับ (Miao et al., 2009: 9172)

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \times 100\% \quad (3-2)$$

$$Precision = \frac{TP}{(TP + FP)} \times 100\% \quad (3-3)$$

$$Recall = \frac{TP}{(TP + FN)} \times 100\% \quad (3-4)$$

3.4.2 การวิเคราะห์ผลกระบวนการคำนวณระดับคะแนนของที่พัก

กระบวนการคำนวณระดับคะแนนของที่พักมีจุดมุ่งหมายเพื่อคำนวณระดับคะแนนของสถานที่พักให้ใกล้เคียงกับคะแนนที่ผู้ใช้ได้ให้ไว้มากที่สุด โดยในการประเมินในส่วนนี้จะใช้ข้อมูล

บทวิจารณ์ที่นำมาทดสอบเกี่ยวกับการประเมินในส่วนแรก (หรือกระบวนการสกัดคุณลักษณะที่สำคัญจากบทวิจารณ์) แต่ใช้ข้อมูลที่เกี่ยวข้องกับเรื่องสถานที่พักเท่านั้น (ทั้งที่สกัดได้ (*TP*) และสกัดไม่ได้ (*FN*)) โดยคุณลักษณะของที่พักที่ระบบสกัดไม่ได้ จะใช้ข้อมูลในชุดทดสอบซึ่งสกัดอย่างถูกต้องไว้แล้ว ดังนั้นในการประเมินผลกระบวนการคำนวณระดับคะแนนนี้ จะได้คุณลักษณะของที่พักทั้งหมดที่อยู่ในบทวิจารณ์ที่นำมาทดสอบ โดยในการประเมินจะใช้วิธีวัดค่าที่เป็นมาตรฐานในการประเมิน ซึ่งในที่นี้จะมี 2 ค่า ได้แก่ ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: *MAE*) และค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: *RMSE*) โดยมีสูตรดังสมการที่ 3-5 และ 3-6 ตามลำดับ (Memmedli and Ozdemi, 2010: 12-21)

$$MAE = \frac{1}{n} \sum_{t=1}^n |A_t - F_t| \quad (3-5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2} \quad (3-6)$$

- เมื่อ A_t คือ ระดับคะแนนที่ผู้ใช้ระบุไว้ในแต่ละบทวิจารณ์ t
 F_t คือ ระดับคะแนนที่ระบบคำนวณได้ในแต่ละบทวิจารณ์ t
 n คือ จำนวนบทวิจารณ์ทั้งหมดที่นำมาทดสอบ

บทที่ 4

ผลการวิจัยและการอภิปรายผล

ในบทนี้กล่าวถึง ผลการพัฒนาระบบ ผลการประเมินความถูกต้องของส่วนอนุমানความรู้ และการอภิปรายผล โดยมีรายละเอียดดังต่อไปนี้

4.1 ผลการพัฒนาระบบ

จากการใช้แนวทางของวงจรการพัฒนา ระบบ ในการพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย ในส่วนของการพัฒนาระบบ ซึ่งเป็นระยะการสร้างระบบงานใหม่ โดยการเขียนโปรแกรมและทดสอบระบบ เพื่อนำโปรแกรมที่ถูกพัฒนาขึ้นไปใช้ในการเก็บรวบรวมบทวิจารณ์ของลูกค้า ผลที่ได้จากการพัฒนาโปรแกรมต้นแบบสำหรับการสกัดบทวิจารณ์และค้นคืนข้อมูลของสถานที่พัก มีรายละเอียดดังแสดงในภาพที่ 4.1



ภาพที่ 4.1 หน้าจอหลักของระบบ

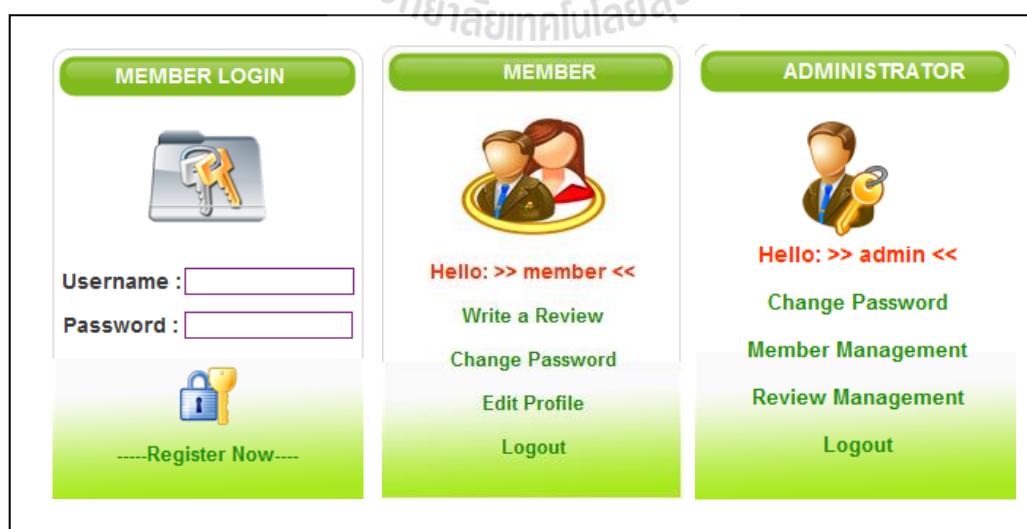
จากภาพที่ 4.1 หน้าหลักของระบบจะประกอบไปด้วยส่วนต่าง ๆ ดังต่อไปนี้

4.1.1 ส่วนการค้นหาข้อมูลที่ปัก ในส่วนนี้ผู้ใช้ทั่วไปสามารถค้นหาที่ปักที่ต้องการทราบข้อมูลได้ โดยการใส่ชื่อบางส่วนของที่ปัก แล้วกดปุ่ม Find!



ภาพที่ 4.2 ส่วนการค้นหาข้อมูลที่ปัก

4.1.2 ส่วนการเข้าสู่ระบบ (Login) ในส่วนนี้ผู้ใช้งานทั่วไปสามารถสมัครสมาชิกเพื่อเขียนบทวิจารณ์ได้ แต่หากเป็นสมาชิกเรียบร้อยแล้วก็สามารถกรอกชื่อผู้ใช้ (Username) และรหัสผ่าน (Password) จากนั้นคลิกที่ปุ่มไอคอนรูปกุญแจเพื่อประมวลผล เมื่อเข้าสู่ระบบสำเร็จแล้ว ระบบจะทำการตรวจสอบสิทธิของผู้ใช้งานว่าเป็นสมาชิก (Member) หรือผู้ดูแลระบบ (Administrator) แล้วจึงแสดงหน้าจอที่มีเมนูการใช้งานตามสิทธิของผู้ใช้งานดังภาพที่ 4.3



ภาพที่ 4.3 ส่วนการเข้าสู่ระบบและเมนูการใช้งานตามสิทธิของผู้ใช้

สิทธิของผู้ใช้งานระบบแบ่งออกเป็น 3 ระดับ ได้แก่ (1) ผู้ใช้งานทั่วไป หมายถึง ผู้ใช้งานที่ยังไม่ได้สมัครเป็นสมาชิกของระบบ ผู้ใช้ประเภทนี้จะสามารถค้นหาและเข้าสู่ข้อมูลแนะนำที่พัก ข้อมูลการส่งเสริมการขายของแหล่งท่องเที่ยว ตลอดจนรูปภาพของที่พักรได้ (2) สมาชิก หมายถึง ผู้ใช้งานที่สมัครเป็นสมาชิกของระบบแล้ว สามารถใช้งานระบบได้เช่นเดียวกับผู้ใช้งานทั่วไป แต่สามารถเขียนบทวิจารณ์ของที่พักร และแก้ไขประวัติส่วนตัวได้ (3) ผู้ดูแลระบบ หมายถึง ผู้ใช้งานที่สมัครเป็นสมาชิกของระบบแล้ว สามารถใช้งานระบบได้เช่นเดียวกับสมาชิก แต่สามารถจัดการข้อมูลของระบบ ข้อมูลบทวิจารณ์ และข้อมูลของสมาชิกได้

4.1.3 ส่วนเมนูหลักในการใช้งานระบบ ในส่วนนี้ผู้ใช้สามารถเข้าสู่การทำงานต่าง ๆ ของระบบได้อย่างรวดเร็ว ประกอบด้วย (1) เมนูหลักสำหรับสมาชิก ได้แก่ การแก้ไขประวัติส่วนตัว และการเขียนบทวิจารณ์ ทั้งนี้ ผู้ใช้ที่จะสามารถใช้งานเมนูหลักสำหรับสมาชิกนี้ได้ จะต้องทำการเข้าสู่ระบบก่อน (2) เมนูหลักสำหรับแสดงการส่งเสริมการขายของที่พักร (3) เมนูหลักสำหรับแนะนำแหล่งท่องเที่ยวตามพื้นที่อำเภอต่าง ๆ



ภาพที่ 4.4 ส่วนเมนูหลักในการใช้งานระบบ

4.1.4 ส่วนแนะนำที่พัก ในส่วนนี้ระบบจะแสดงรายละเอียดเบื้องต้น พร้อมทั้งระดับความพึงพอใจโดยรวมของที่พักรนั้นตั้งแต่ระดับ 1-5 ดาว ผู้ใช้งานที่สนใจ สามารถคลิกเพื่ออ่านรายละเอียดตลอดจนบทวิจารณ์ของที่พักรนั้น ๆ ได้

Accommodation Recommendation



Belle Villa Resort
Khaoyai



Villa Paradis Hotel



Kratomhin Nuntapuk
HotelResort



Sima Thani Hotel

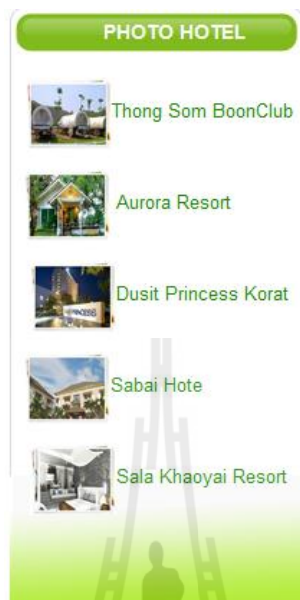
ภาพที่ 4.5 ส่วนแนะนำที่พัก

4.1.5 ส่วนแสดงการส่งเสริมการขายของที่พัก ในส่วนนี้จะแสดงข้อมูลการส่งเสริมการขายของที่พักต่าง ๆ ผู้ใช้งานที่สนใจสามารถคลิกเพื่ออ่านรายละเอียดของแต่ละที่พักได้



ภาพที่ 4.6 ส่วนแสดงการส่งเสริมการขายของที่พัก

4.1.6 ส่วนแสดงรูปภาพของที่พักร ในส่วนนี้ผู้ใช้งานสามารถชมรูปภาพของแต่ละที่พัก เพื่อประกอบการตัดสินใจไปท่องเที่ยวได้

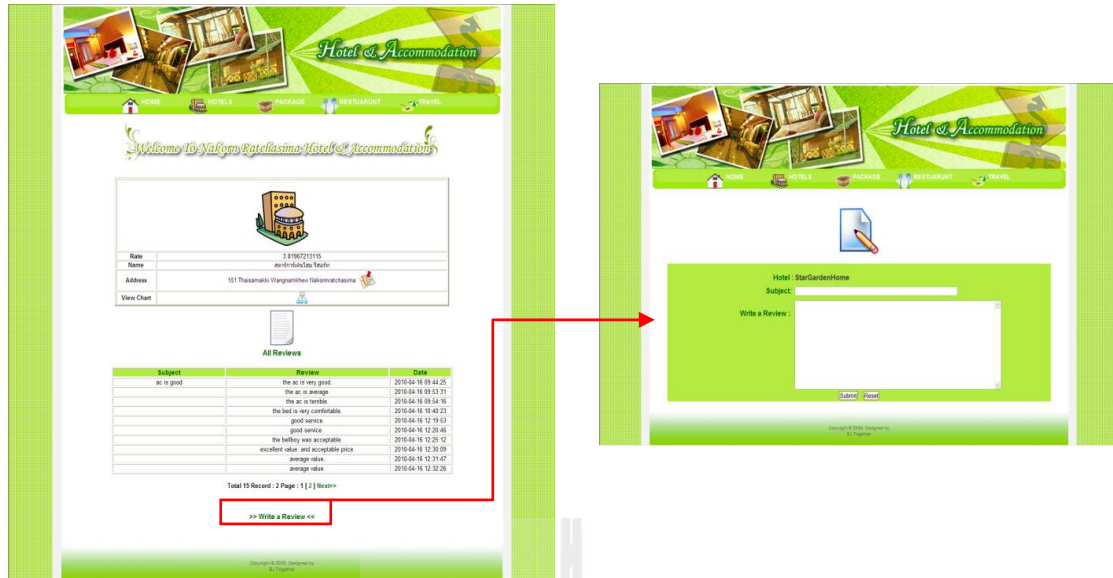


ภาพที่ 4.7 ส่วนแสดงรูปภาพของที่พักร

เมื่อผู้ใช้งานต้องการเขียนบทวิจารณ์เพื่อแสดงความคิดเห็นเกี่ยวกับที่พักที่ได้ไปมาแล้ว จะต้องปฏิบัติตามขั้นตอนการใช้งาน ดังนี้

1. ผู้ใช้งานทำการเข้าสู่ระบบ ในส่วนของการเข้าสู่ระบบ
2. ผู้ใช้งานทำการค้นหาที่พักที่ต้องการเขียนบทวิจารณ์ โดยสามารถทำได้ 3 วิธีดังต่อไปนี้

วิธีที่ 1 ใส่ชื่อที่พักที่ต้องการค้นหา ในส่วนการค้นหาข้อมูลที่พักร เมื่อค้นหาพบแล้ว ผู้ใช้สามารถคลิกดูรายละเอียด บทวิจารณ์ และสามารถเขียนบทวิจารณ์เพิ่มสำหรับที่พักนั้นได้ ดังแสดงในภาพที่ 4.8



ภาพที่ 4.8 ส่วนแสดงการค้นหาที่พักและเขียนบทวิจารณ์ วิธีที่ 1

วิธีที่ 2 กรณีมีสิทธิการใช้งานเป็นสมาชิก ผู้ใช้งานสามารถคลิกที่เมนูการใช้งาน “Write a Review” ในส่วนการเข้าสู่ระบบเพื่อค้นหาที่พักที่ต้องการ เมื่อค้นหาพบแล้ว ผู้ใช้สามารถคลิกดูบทวิจารณ์ แผนภาพต้นไม้การจัดระดับความพึงพอใจของที่พัก และสามารถเขียนบทวิจารณ์เพิ่มสำหรับที่พักนั้นได้ ดังแสดงในภาพที่ 4.9

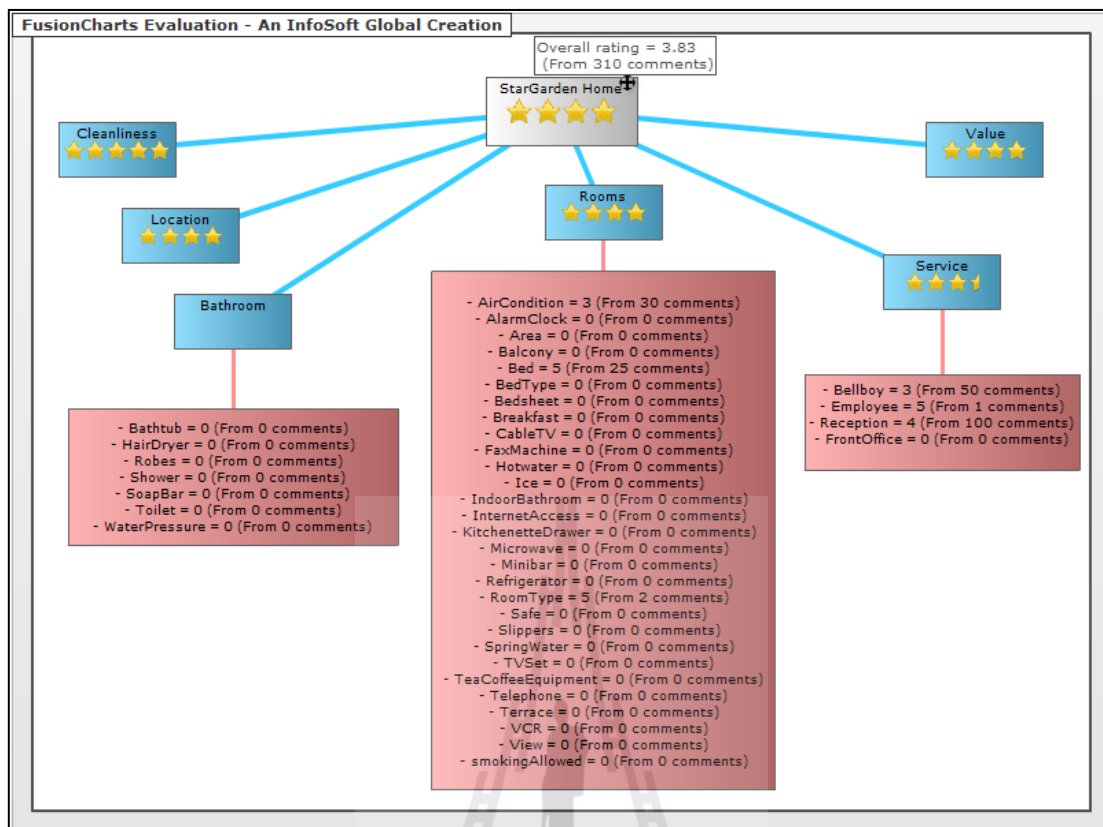


ภาพที่ 4.9 ส่วนแสดงการค้นหาที่พักและเขียนบทวิจารณ์ วิธีที่ 2

วิธีที่ 3 กรณีมีสิทธิการใช้งานเป็นสมาชิกหรือผู้ดูแลระบบ ผู้ใช้งานสามารถคลิกที่เมนูการใช้งาน “Write Review” ในส่วนเมนูสำหรับสมาชิกเพื่อค้นหาที่พักที่ต้องการ จากนั้นจึงปฏิบัติเช่นเดียวกับวิธีที่ 2

3. เมื่อผู้ใช้งานเขียนบทวิจารณ์เรียบร้อยแล้ว ให้คลิกที่ปุ่ม “Submit” จากนั้นระบบจะทำการบันทึกข้อมูลและประมวลผลการสกัดความรู้ พร้อมทั้งแสดงผลภาพจากการสกัดความรู้ว่า บทวิจารณ์ที่รับเข้ามามีคำวิจารณ์เกี่ยวข้องกับเรื่องใดบ้างและคิดเป็นระดับคะแนนเท่าใด หากบทวิจารณ์นั้นมีคำวิจารณ์หลายเรื่อง ระบบจะทำการประมวลผลทีละเรื่อง ผู้ใช้สามารถเลือกที่จะแสดงแผนภาพต้นไม้การจัดระดับความพึงพอใจของที่พักนั้นทันที ดังแสดงในภาพที่ 4.10 โดยคลิกที่จุดเชื่อมโยง “View Chart” หรือประมวลผลคำวิจารณ์เรื่องถัดไป โดยคลิกที่จุดเชื่อมโยง “Analyze now !!” ดังแสดงในภาพที่ 4.11 หากคำวิจารณ์เรื่องใดระบบไม่สามารถสกัดความรู้ได้ จะแสดงข้อความ “Cannot analyze Total Score of this sentence.” พร้อมทั้งแสดงจุดเชื่อมโยงให้เขียนบทความอีกครั้ง ดังแสดงในภาพที่ 4.12





ภาพที่ 4.10 แผนภาพต้นไม้การจัดระดับความพึงพอใจของที่พักร

HOME HOTELS PACKAGE RESTUARUNT TRAVEL

Score of this comment = 4

>> View Chart of "StarGardenHome" <<

(But there are next sentence remaining. >> Analyze now !! <<)

Copyright © 2009. Designed by BJ Together

ภาพที่ 4.11 ส่วนแสดงผลการสกัดความรู้จากคำวิจารณ์ ในกรณีมีคำวิจารณ์หลายเรื่อง



ภาพที่ 4.12 ผลการสกัดความรู้จากคำวิจารณ์ ในกรณีที่ระบบไม่สามารถสกัดได้

4.1.7 ส่วนแสดงสารสนเทศทางภูมิศาสตร์ออนไลน์

ส่วนแสดงสารสนเทศทางภูมิศาสตร์จะถูกนำมาใช้เพื่อสนับสนุนการตัดสินใจของผู้ใช้ ในการเลือกสถานที่พักเพิ่มเติมจากระดับคะแนนที่วิเคราะห์มาได้ โดยแสดงสิ่งอำนวยความสะดวก บริเวณรอบของสถานที่พักเหล่านั้น ดังนี้

1) บริการสาธารณะ ประกอบด้วย ปิมน้ำมัน ปิมน้ำมัน (NGV) ปิมนแอลพีจี (LPG) ตู้เอทีเอ็ม ธนาคาร และศูนย์บริการการท่องเที่ยว โดยในภาพที่ 4.13 คือการแสดงผลเมื่อคลิกเลือก ให้บริการสาธารณะทั้งหมดแสดงในแผนที่



ภาพที่ 4.13 หน้าจอส่วนแสดงสารสนเทศทางภูมิศาสตร์ออนไลน์ของบริการสาธารณะรอบสถานที่พัก

เป็นข้อมูลคำวิจารณ์ที่เกี่ยวข้องจำนวน 205 คำวิจารณ์ ข้อมูลคำวิจารณ์ที่ไม่เกี่ยวข้องจำนวน 313 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ผิดไวยากรณ์จำนวน 143 คำวิจารณ์ รวมทั้งสิ้น 661 คำวิจารณ์

ในการทดสอบความรู้ที่สกัดได้จากระบบ ผลการทดสอบจะพิจารณาถึงค่ามาตรฐานทั่วไปที่ใช้ในการประเมินความถูกต้องของการสกัดสารสนเทศ รวมทั้งพิจารณาถึงระดับคะแนนในภาพรวมของสถานที่พักที่ระบบสกัดได้ โดยการทดสอบแบ่งออกเป็น 3 ประเด็น ได้แก่

ประเด็นที่ 1 การทดสอบความถูกต้องในการสกัดคุณลักษณะที่สำคัญของสถานที่พักจากบทวิจารณ์ออนไลน์แบบอัตโนมัติ โดยใช้วิธีวัดค่ามาตรฐานในการประเมินความถูกต้อง 3 ค่า ได้แก่ ค่าความถูกต้อง (*Accuracy*) ค่าความแม่นยำ (*Precision*) และค่าความระลึก (*Recall*) ซึ่งวัตถุประสงค์ของการทดสอบในประเด็นนี้คือเพื่อประเมินความถูกต้องของขั้นตอนวิธีการสกัดคุณลักษณะที่สำคัญ (*Feature Extraction*) ที่ได้พัฒนาขึ้นมา

ประเด็นที่ 2 การทดสอบเพื่อหาค่าระดับคะแนนความสัมพันธ์ระหว่างคำวิเศษณ์และคำคุณศัพท์ (*r*) ที่เหมาะสมกับด้านการท่องเที่ยว โดยในการทดสอบจะทำโดยการเปรียบเทียบค่าระดับคะแนนที่ระบบคำนวณได้กับค่าที่นักท่องเที่ยวกำหนด ค่าระดับความสัมพันธ์ที่เหมาะสมคือ ค่าที่ทำให้ค่าความคลาดเคลื่อนสัมบูรณ์มีค่าต่ำที่สุด โดยวิธีการที่ใช้ในการคำนวณระดับคะแนนของระบบในที่นี่จะทดสอบ 2 วิธี คือ วิธีที่ใช้ตรรกศาสตร์คลุมเครือ และวิธีการที่ไม่ได้ใช้ตรรกศาสตร์คลุมเครือ

ประเด็นที่ 3 การทดสอบความใกล้เคียงของระดับคะแนนในภาพรวมของทั้งโรงแรมที่ระบบสกัดได้กับที่นักท่องเที่ยวกำหนด ทั้งในกรณีที่ใช้ตรรกศาสตร์คลุมเครือและไม่ใช้ตรรกศาสตร์คลุมเครือในการคำนวณระดับคะแนน โดยใช้วิธีวัดค่ามาตรฐานในการประเมินความผิดพลาด 2 ค่า ได้แก่ ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (*MAE*) และค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (*RMSE*) ซึ่งวัตถุประสงค์ของการทดสอบในประเด็นนี้คือเพื่อเปรียบเทียบความใกล้เคียงของระดับคะแนนในภาพรวมของทั้งโรงแรมที่ระบบคำนวณได้ โดยใช้ระดับคะแนนที่นักท่องเที่ยวกำหนดไว้เป็นเป้าหมายในการวัดความถูกต้องของระดับคะแนน

สำหรับกระบวนการคำนวณระดับคะแนนในภาพรวมของระบบโดยใช้เทคนิคตรรกศาสตร์คลุมเครือจะดำเนินการดังที่ได้กล่าวมาแล้วในบทที่ 3 ในส่วนอนุมานความรู้ แต่สำหรับกระบวนการคำนวณระดับคะแนนในภาพรวมของระบบโดยไม่ใช้ตรรกศาสตร์คลุมเครือจะใช้วิธีการหาค่าเฉลี่ยของระดับคะแนนเฉพาะเรื่องทุกอย่างที่นักท่องเที่ยวกำหนด ดังสมการที่ 4.1

$$\text{Overall Score} = \frac{\sum_{i=1}^n \text{Score}_i}{n} \quad (4.1)$$

โดยที่ *Overall Score* หมายถึง ระดับคะแนนในภาพรวมของทั้งโรงแรม; *Score_i* หมายถึง ระดับคะแนนในเรื่องวิจารณ์ที่ *i*; และ *n* หมายถึง จำนวนเรื่องทั้งหมดที่วิจารณ์

จากการจำลองสภาพแวดล้อมเพื่อทดสอบระบบ สามารถแยกเป็น 3 ประเด็นดังที่กล่าวมาแล้ว โดยมีรายละเอียดของผลการประเมิน ดังต่อไปนี้

4.2.1 ผลการประเมินความถูกต้องของการสกัดคุณลักษณะที่สำคัญของสถานที่พัก

ในการวัดความถูกต้องของการสกัดคุณลักษณะที่สำคัญของสถานที่พัก (Feature Extraction) ได้จำแนกข้อมูลผลการทดสอบตามระดับความพึงพอใจที่ต่างกัน 5 ระดับ ดังต่อไปนี้

- **ระดับความพึงพอใจที่ 1** สามารถจำแนกข้อมูลผลการทดสอบออกเป็น 4 กลุ่ม ดังปรากฏในตารางที่ 4.1 ซึ่งแสดงว่าข้อมูลคำวิจารณ์ที่สามารถสกัดได้และมีความเกี่ยวข้อง (TP) มีจำนวน 21 คำวิจารณ์; ไม่มีข้อมูลคำวิจารณ์ที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (FP); ข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (TN) มีจำนวน 78 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (FN) มีจำนวน 13 คำวิจารณ์ จากค่าของผลการทดสอบดังกล่าวสามารถคำนวณค่าความถูกต้องได้เท่ากับ 88.39% ค่าความแม่นยำได้เท่ากับ 100% และค่าความระลึกลับเท่ากับ 61.76% ดังแสดงในสมการที่ 4-1, 4-2 และ 4-3 ตามลำดับ

ตารางที่ 4.1 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 1

	ข้อมูลที่มีความเกี่ยวข้อง (+)	ข้อมูลที่ไม่มีความเกี่ยวข้อง (-)
ข้อมูลที่สกัดได้ (+)	TP = 21	FP = 0
ข้อมูลที่สกัดไม่ได้ (-)	FN = 13	TN = 78

$$Accuracy = \frac{21+78}{(21+78+0+13)} \times 100\% = 88.39\% \quad (4-1)$$

$$Precision = \frac{21}{(21+0)} \times 100\% = 100\% \quad (4-2)$$

$$Recall = \frac{21}{(21+13)} \times 100\% = 61.76\% \quad (4-3)$$

- **ระดับความพึงพอใจที่ 2** สามารถจำแนกข้อมูลผลการทดสอบออกเป็น 4 กลุ่ม ดังปรากฏในตารางที่ 4.2 ซึ่งแสดงว่าข้อมูลคำวิจารณ์ที่สามารถสกัดได้และมีความเกี่ยวข้อง (TP) มีจำนวน 41 คำวิจารณ์; ไม่มีข้อมูลคำวิจารณ์ที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (FP); ข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (TN) มีจำนวน 83 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ไม่สามารถ

สกัดได้แต่มีความเกี่ยวข้อง (FN) มีจำนวน 8 คำวิจารณ์ จากค่าของผลการทดสอบดังกล่าวสามารถคำนวณค่าความถูกต้องได้เท่ากับ 93.94% ค่าความแม่นยำได้เท่ากับ 100% และค่าความระลึกลับเท่ากับ 83.67% ดังแสดงในสมการที่ 4-4, 4-5 และ 4-6 ตามลำดับ

ตารางที่ 4.2 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 2

	ข้อมูลที่มีความเกี่ยวข้อง (+)	ข้อมูลที่ไม่มีความเกี่ยวข้อง (-)
ข้อมูลที่สกัดได้ (+)	TP = 41	FP = 0
ข้อมูลที่สกัดไม่ได้ (-)	FN = 8	TN = 83

$$Accuracy = \frac{41+83}{(41+83+0+8)} \times 100\% = 93.94\% \quad (4-4)$$

$$Precision = \frac{41}{(41+0)} \times 100\% = 100\% \quad (4-5)$$

$$Recall = \frac{41}{(41+8)} \times 100\% = 83.67\% \quad (4-6)$$

- ระดับความพึงพอใจที่ 3 สามารถจำแนกข้อมูลผลการทดสอบออกเป็น 4 กลุ่ม ดังปรากฏในตารางที่ 4.3 ซึ่งแสดงว่าข้อมูลคำวิจารณ์ที่สามารถสกัดได้และมีความเกี่ยวข้อง (TP) มีจำนวน 34 คำวิจารณ์; ไม่มีข้อมูลคำวิจารณ์ที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (FP); ข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (TN) มีจำนวน 46 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (FN) มีจำนวน 11 คำวิจารณ์ จากค่าของผลการทดสอบดังกล่าวสามารถคำนวณค่าความถูกต้องได้เท่ากับ 87.91% ค่าความแม่นยำได้เท่ากับ 100% และค่าความระลึกลับเท่ากับ 75.56% ดังแสดงในสมการที่ 4-7, 4-8 และ 4-9 ตามลำดับ

ตารางที่ 4.3 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 3

	ข้อมูลที่มีความเกี่ยวข้อง (+)	ข้อมูลที่ไม่มีความเกี่ยวข้อง (-)
ข้อมูลที่สกัดได้ (+)	TP = 34	FP = 0
ข้อมูลที่สกัดไม่ได้ (-)	FN = 11	TN = 46

$$Accuracy = \frac{34+46}{(34+46+0+11)} \times 100\% = 87.91\% \quad (4-7)$$

$$Precision = \frac{34}{(34+0)} \times 100\% = 100\% \quad (4-8)$$

$$Recall = \frac{34}{(34+11)} \times 100\% = 75.56\% \quad (4-9)$$

- **ระดับความพึงพอใจที่ 4** สามารถจำแนกข้อมูลผลการทดสอบออกเป็น 4 กลุ่ม ดังปรากฏในตารางที่ 4.4 ซึ่งแสดงว่าข้อมูลคำวิจารณ์ที่สามารถสกัดได้และมีความเกี่ยวข้อง (TP) มีจำนวน 38 คำวิจารณ์; ไม่มีข้อมูลคำวิจารณ์ที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (FP); ข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (TN) มีจำนวน 52 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (FN) มีจำนวน 15 คำวิจารณ์ จากค่าของผลการทดสอบดังกล่าวสามารถคำนวณค่าความถูกต้องได้เท่ากับ 85.71% ค่าความแม่นยำได้เท่ากับ 100% และค่าความระลึกเท่ากับ 71.70% ดังแสดงในสมการที่ 4-10, 4-11 และ 4-12 ตามลำดับ

ตารางที่ 4.4 จำนวนข้อมูลที่นำมาใช้ในการประเมินผล ระดับความพึงพอใจที่ 4

	ข้อมูลที่มีความเกี่ยวข้อง (+)	ข้อมูลที่ไม่มีความเกี่ยวข้อง (-)
ข้อมูลที่สกัดได้ (+)	TP = 38	FP = 0
ข้อมูลที่สกัดไม่ได้ (-)	FN = 15	TN = 52

$$Accuracy = \frac{38+52}{(38+52+0+15)} \times 100\% = 85.71\% \quad (4-10)$$

$$Precision = \frac{38}{(38+0)} \times 100\% = 100\% \quad (4-11)$$

$$Recall = \frac{38}{(38+15)} \times 100\% = 71.70\% \quad (4-12)$$

- **ระดับความพึงพอใจที่ 5** สามารถจำแนกข้อมูลผลการทดสอบออกเป็น 4 กลุ่ม ดังปรากฏในตารางที่ 4.5 ซึ่งแสดงว่าข้อมูลคำวิจารณ์ที่สามารถสกัดได้และมีความเกี่ยวข้อง (TP) มีจำนวน 20 คำวิจารณ์; ไม่มีข้อมูลคำวิจารณ์ที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (FP); ข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (TN) มีจำนวน 54 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (FN) มีจำนวน 4 คำวิจารณ์ จากค่าของผลการทดสอบดังกล่าวสามารถ

คำนวณค่าความถูกต้องได้เท่ากับ 94.87% ค่าความแม่นยำได้เท่ากับ 100% และค่าความระลึกลับเท่ากับ 83.33% ดังแสดงในสมการที่ 4-13, 4-14 และ 4-15 ตามลำดับ

ตารางที่ 4.5 จำนวนข้อมูลที่ใช้ในการประเมินผล ระดับความพึงพอใจที่ 5

	ข้อมูลที่มีความเกี่ยวข้อง (+)	ข้อมูลที่ไม่มีความเกี่ยวข้อง (-)
ข้อมูลที่สกัดได้ (+)	TP = 20	FP = 0
ข้อมูลที่สกัดไม่ได้ (-)	FN = 4	TN = 54

$$Accuracy = \frac{20+54}{(20+54+0+4)} \times 100\% = 94.87\% \quad (4-13)$$

$$Precision = \frac{20}{(20+0)} \times 100\% = 100\% \quad (4-14)$$

$$Recall = \frac{20}{(20+4)} \times 100\% = 83.33\% \quad (4-15)$$

- **ระดับความพึงพอใจโดยรวม** สามารถจำแนกข้อมูลผลการทดสอบออกเป็น 4 กลุ่ม ดังปรากฏในตารางที่ 4.6 ซึ่งแสดงว่าข้อมูลคำวิจารณ์ที่สามารถสกัดได้และมีความเกี่ยวข้อง (TP) มีจำนวน 154 คำวิจารณ์; ไม่มีข้อมูลคำวิจารณ์ที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (FP); ข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (TN) มีจำนวน 313 คำวิจารณ์ และข้อมูลคำวิจารณ์ที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (FN) มีจำนวน 51 คำวิจารณ์ จากค่าของผลการทดสอบดังกล่าวสามารถคำนวณค่าความถูกต้องได้เท่ากับ 90.15% ค่าความแม่นยำได้เท่ากับ 100% และค่าความระลึกลับเท่ากับ 75.12% ดังแสดงในสมการที่ 4-16, 4-17 และ 4-18 ตามลำดับ

ตารางที่ 4.6 จำนวนข้อมูลที่ใช้ในการประเมินผล ระดับความพึงพอใจโดยรวม

	ข้อมูลที่มีความเกี่ยวข้อง (+)	ข้อมูลที่ไม่มีความเกี่ยวข้อง (-)
ข้อมูลที่สกัดได้ (+)	TP = 154	FP = 0
ข้อมูลที่สกัดไม่ได้ (-)	FN = 51	TN = 313

$$Accuracy = \frac{154+313}{(154+313+0+51)} \times 100\% = 90.15\% \quad (4-16)$$

$$\text{Precision} = \frac{154}{(154+0)} \times 100\% = 100\% \quad (4-17)$$

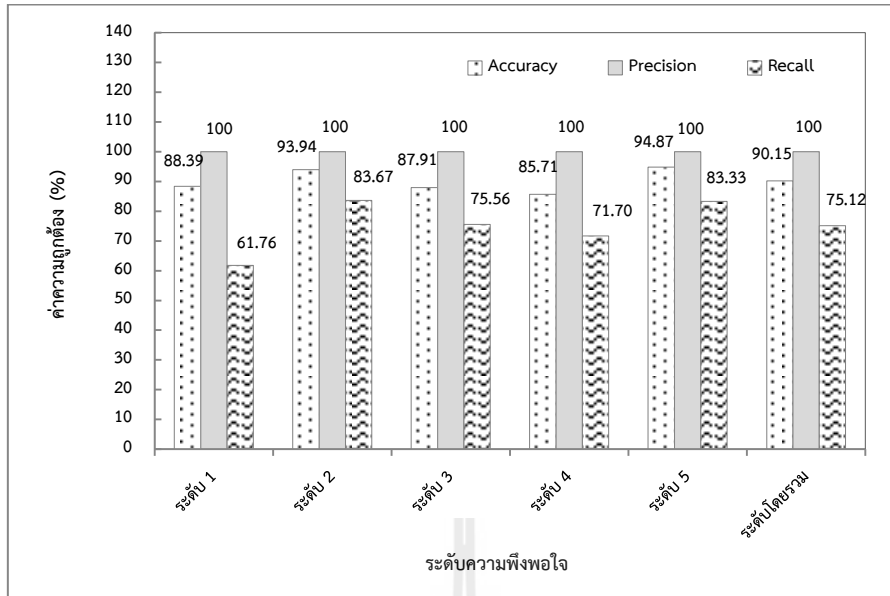
$$\text{Recall} = \frac{154}{(154+51)} \times 100\% = 75.12\% \quad (4-18)$$

อย่างไรก็ตาม ระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทยดังกล่าว จะก่อให้เกิดประโยชน์สูงสุดก็ต่อเมื่อผู้ใช้งานมีการใช้ภาษาที่ถูกต้องตามหลักไวยากรณ์ภาษาอังกฤษ และเขียนประโยคคำวิจารณ์ที่ไม่ยาวจนเกินไป เพื่อนำไปประมวลผลลัพธ์ที่ถูกต้องและสอดคล้องกับการตีความหมายของผู้ใช้งานมากยิ่งขึ้น จากการประเมินความถูกต้องของการสกัดคุณลักษณะที่สำคัญของสถานที่พัก (Feature Extraction) สามารถวิเคราะห์และอภิปรายผลได้ดังต่อไปนี้

1. เมื่อพิจารณาค่าความแม่นยำ (Precision) ของทุกระดับความพึงพอใจ พบว่ามีค่าเท่ากับ 100% เสมอ เนื่องจากการศึกษาครั้งนี้ได้สกัดความรู้โดยใช้ออนไลน์ที่มีจำนวนคำศัพท์คงที่และมีความเกี่ยวข้องเฉพาะกับเรื่องที่พักเท่านั้น ดังนั้น ระบบจะไม่สามารถสกัดข้อมูลที่ไม่มีความเกี่ยวข้องกับเรื่องที่พักได้ ($FP = 0$) จึงทำให้ค่าความแม่นยำที่คำนวณได้มีค่าเท่ากับ 100% เสมอ ดังสูตร

$$\text{precision} = \frac{TP}{(TP+FP)} \times 100\% = \frac{TP}{(TP+0)} \times 100\% = 100\%$$

2. เมื่อพิจารณาค่าความระลึก (Recall) ของการสกัดความรู้ พบว่าระดับความพึงพอใจที่ 2 มีค่ามากที่สุด (83.67%) รองลงมาคือ ระดับความพึงพอใจที่ 5 (83.33%), ระดับความพึงพอใจที่ 3 (75.56%), ระดับความพึงพอใจที่ 4 (71.70%) และระดับความพึงพอใจที่ 1 ซึ่งมีค่าน้อยที่สุด (61.76%) ตามลำดับ โดยมีค่าความระลึกของการสกัดความรู้โดยรวมคิดเป็น 75.12% ดังแสดงในภาพที่ 4.18



ภาพที่ 4.18 แผนภูมิสรุปผลการประเมินความถูกต้องของการสกัดคุณลักษณะที่สำคัญของสถานที่พัก

สำหรับสาเหตุที่ทำให้ระดับความพึงพอใจที่ 1 มีค่าอยู่ในเกณฑ์ต่ำกว่าระดับความพึงพอใจอื่น อันเนื่องมาจากกลุ่มของกฎไวยากรณ์ไม่พึงบริบทยังไม่ครอบคลุมรูปแบบประโยคคำวิจารณ์ส่วนใหญ่ที่ผู้ใช้ทำการวิจารณ์ เพราะในบทวิจารณ์ที่แสดงถึงความไม่พอใจในระดับที่ 1 นั้น ผู้ใช้ส่วนใหญ่จะใช้ประโยคคำวิจารณ์ในลักษณะบรรยายมากกว่าการกล่าวถึงเรื่องที่ไม่พอใจโดยตรง ทำให้กระบวนการสกัดความรู้ที่ทำที่ละประโยคดังที่นำเสนอในงานวิจัยนี้ ไม่สามารถสกัดได้อย่างถูกต้อง เมื่อเปรียบเทียบกับบทวิจารณ์ในระดับความพึงพอใจอื่น ๆ ที่ผู้ใช้มักจะวิจารณ์อย่างตรงไปตรงมา การสกัดที่ละประโยคจึงมีความถูกต้อง ทำให้ค่าความระลึกที่ได้สูงกว่าระดับความพึงพอใจที่ 1

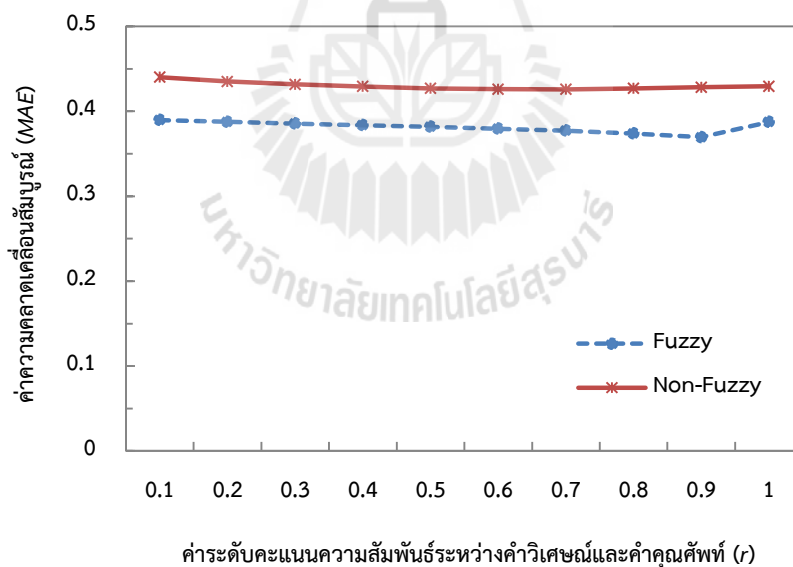
3. เมื่อพิจารณาค่าความถูกต้อง (Accuracy) ของการสกัดความรู้ พบว่าระดับความพึงพอใจที่ 5 มีค่ามากที่สุด (94.87%) รองลงมาคือ ระดับความพึงพอใจที่ 2 (93.94%), ระดับความพึงพอใจที่ 1 (88.39%), ระดับความพึงพอใจที่ 3 (87.91%) และระดับความพึงพอใจที่ 4 ซึ่งมีค่าน้อยที่สุด (85.71%) ตามลำดับ โดยมีค่าความถูกต้องของการสกัดความรู้โดยรวมคิดเป็น 90.15% ดังแสดงในภาพที่ 4.18 เหตุผลที่ทำให้ค่าดังกล่าวอยู่ในเกณฑ์ค่อนข้างสูง เนื่องจากการวิจารณ์ถึงเรื่องที่ไม่เกี่ยวข้องกับขอบเขตที่ระบบสนใจ มีผลทำให้การประเมินความถูกต้องในการสกัดความรู้ของระบบมีค่าข้อมูลที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (*True Negative: TN*) เป็นจำนวนมาก จากค่าความถูกต้องที่เกิดจากสูตร
$$accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \times 100\%$$
 เมื่อนำค่าดังกล่าวมารวมกับค่า *TP* จะทำให้ผลรวมที่ได้มีค่าสูงกว่าข้อมูลที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (*False Negative: FN*) เป็นจำนวนมากจนทำให้ค่า *FN* ส่งผลต่อการคำนวณน้อยหรืออาจไม่ส่งผลเลย นอกจากนี้ ค่า *FP* ยังมี

ค่าเท่ากับ 0 เสมอตั้งที่กล่าวมาแล้วในหัวข้อที่ 1 จึงทำให้ค่าความถูกต้องของการสกัดความรู้ที่คำนวณได้สูงกว่าค่าความระลึกลในทุกระดับ เมื่อเทียบกับการคำนวณค่าความระลึกลที่มาจากสูตร

$$recall = \frac{TP}{(TP + FN)} \times 100\%$$

4.2.2 ผลการหาค่าระดับคะแนนความสัมพันธ์ระหว่างค่าวิเศษณ์และค่าคุณศัพท์

งานวิจัยนี้ได้นำเสนอผลการวิเคราะห์สำหรับการหาค่าระดับคะแนนความสัมพันธ์ระหว่างค่าวิเศษณ์และค่าคุณศัพท์ (r) ที่เหมาะสมกับด้านการท่องเที่ยว ดังภาพที่ 4.19 โดยในการทดสอบจะทำโดยการเปรียบเทียบค่าระดับคะแนนที่ระบบคำนวณได้กับค่าที่ผู้ใช้กำหนด ค่าระดับความสัมพันธ์ที่เหมาะสมคือ ค่าที่ทำให้ค่าความคลาดเคลื่อนสัมบูรณ์ (MAE) มีค่าต่ำที่สุด โดยวิธีการที่ใช้ในการคำนวณระดับของระบบในที่นี้ จะมี 2 วิธี คือ วิธีที่ใช้ตรรกศาสตร์คลุมเครือและวิธีการที่ไม่ได้ใช้ตรรกศาสตร์คลุมเครือ ดังนั้น ในการทดสอบนี้จะใช้เพื่อหาค่าระดับคะแนนความสัมพันธ์ระหว่างค่าวิเศษณ์และค่าคุณศัพท์ (r) ที่เหมาะสมของแต่ละวิธีนั้น ซึ่งจากภาพที่ 4.19 จะเห็นได้ว่า ค่า $r=0.9$ จะได้รับค่าความคลาดเคลื่อนสัมบูรณ์ที่ต่ำที่สุดสำหรับวิธีการที่ใช้ตรรกศาสตร์คลุมเครือ และค่า $r=0.7$ จะได้รับค่าความคลาดเคลื่อนสัมบูรณ์ที่ต่ำที่สุดสำหรับวิธีการที่ไม่ได้ใช้ตรรกศาสตร์คลุมเครือ



ภาพที่ 4.19 ผลการวิเคราะห์ค่าระดับคะแนนความสัมพันธ์ระหว่างค่าวิเศษณ์และค่าคุณศัพท์ (r)

4.2.3 ผลการประเมินความถูกต้องของการคำนวณระดับคะแนนของสถานที่พัก

ในการศึกษานี้ได้ใช้ข้อมูลบทวิจารณ์ จำนวน 100 บทวิจารณ์ ซึ่งอยู่ในระดับความพึงพอใจที่ต่างกัน 5 ระดับ จำนวนระดับละ 20 บทวิจารณ์ เช่นเดียวกับการประเมินความถูกต้องของการสกัดคุณลักษณะที่สำคัญของสถานที่พัก แต่จะสกัดคุณลักษณะที่สำคัญของสถานที่พักทั้งหมดที่มีอยู่ในบทวิจารณ์มาใช้เป็นกรณีทดสอบ (Test Case) นั่นก็คือ ไม่มีคุณลักษณะที่สำคัญของสถานที่พัก แต่สกัดไม่ได้ (หรือค่า *FN*) มาใช้ในการประเมินกระบวนการคำนวณระดับคะแนน

สำหรับการประเมินความถูกต้องของกระบวนการคำนวณระดับคะแนนของระบบทำโดยการทดสอบความใกล้เคียงกับระดับคะแนนที่นักท่องเที่ยวกำหนด โดยวิธีการในการคำนวณระดับคะแนนของระบบมี 2 วิธี คือ วิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือ (Fuzzy) และวิธีการซึ่งไม่ใช้ตรรกศาสตร์คลุมเครือ (Non-fuzzy) โดยทั้งสองวิธีได้ใช้วิธีวัดค่ามาตรฐานในการประเมินความผิดพลาด 2 ค่า ได้แก่ ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAE) และค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE)

ในตารางที่ 4.7 แสดงค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) และค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ของการให้คะแนนภาพรวม โดยเปรียบเทียบวิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือ และวิธีการซึ่งไม่ใช้ตรรกศาสตร์คลุมเครือ จากผลการทดลอง แสดงให้เห็นว่า ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ของทั้งสองวิธีมีค่าใกล้เคียงศูนย์ ซึ่งอธิบายได้ว่า ในทุกกรณีทดสอบ (Test Case) ของแต่ละระดับความพึงพอใจมีค่าคลาดเคลื่อนน้อยมาก ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) ของวิธีการที่ใช้ตรรกศาสตร์คลุมเครือมีค่าน้อยกว่าวิธีการที่ไม่ใช้ตรรกศาสตร์คลุมเครือเล็กน้อย

ตารางที่ 4.7 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) และค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ระหว่างวิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือและวิธีการซึ่งไม่ใช้ตรรกศาสตร์คลุมเครือ

กรณีทดสอบ (Test Case)	วิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือ (Fuzzy)		วิธีการซึ่งไม่ได้ใช้ตรรกศาสตร์คลุมเครือ (Non-fuzzy)	
	MAE	RMSE	MAE	RMSE
ระดับความพึงพอใจ 1	0.428	0.470	0.513	0.589
ระดับความพึงพอใจ 2	0.431	0.555	0.548	0.635
ระดับความพึงพอใจ 3	0.302	0.358	0.331	0.381
ระดับความพึงพอใจ 4	0.332	0.413	0.344	0.436
ระดับความพึงพอใจ 5	0.368	0.511	0.415	0.548
<i>ภาพรวม</i>	<i>0.370</i>	<i>0.466</i>	<i>0.426</i>	<i>0.523</i>

อย่างไรก็ตาม ถึงแม้ว่าภาพรวมของค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ของวิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือ กับวิธีการซึ่งไม่ใช้ตรรกศาสตร์คลุมเครือ อาจมีความแตกต่างอย่างไม่มีนัยสำคัญ แต่จำนวนของกรณีทดสอบของวิธีการที่ใช้ตรรกศาสตร์คลุมเครือ ซึ่งมีค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์น้อยกว่า 0.5 คะแนน (MAE < 0.5) มีจำนวนมากกว่าวิธีการซึ่งไม่ใช้ตรรกศาสตร์คลุมเครืออย่างเห็นได้ชัดในทุก ๆ ระดับความพึงพอใจ ดังแสดงในตารางที่ 4.8 ซึ่งจากตารางจะเห็นว่า จำนวนกรณีทดสอบซึ่งมีค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์น้อยกว่า 0.5 คะแนน ของวิธีการที่ใช้ตรรกศาสตร์คลุมเครือมีสูงถึงร้อยละ 70.16 แต่วิธีการซึ่งไม่ใช้ตรรกศาสตร์คลุมเครือมีเพียงร้อยละ 56.68 ซึ่งกล่าวได้ว่า กระบวนการคำนวณระดับคะแนนของสถานที่พักซึ่งใช้ตรรกศาสตร์คลุมเครือมีความถูกต้องหรือใกล้เคียงกับระดับความพึงพอใจของนักท่องเที่ยวมากกว่าวิธีการซึ่งไม่ใช้ตรรกศาสตร์คลุมเครือ

ซึ่งเหตุผลที่ทำให้การคำนวณระดับคะแนนโดยระบบ แตกต่างกับการให้คะแนนโดยนักท่องเที่ยวกำหนดก็คือ เกณฑ์การให้คะแนนของแต่ละค่าในฐานความรู้ซึ่งกำหนดโดยผู้เชี่ยวชาญด้านภาษา อาจไม่สอดคล้องกับความคิดเห็นของนักท่องเที่ยว

ตารางที่ 4.8 การเปรียบเทียบร้อยละของค่าผิดพลาดระหว่างวิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือและวิธีการซึ่งไม่ได้ใช้ตรรกศาสตร์คลุมเครือ

กรณีทดสอบ (Test Case)	ร้อยละของค่าผิดพลาด ด้วยวิธีการซึ่งใช้ตรรกศาสตร์คลุมเครือ (Fuzzy)		ร้อยละของค่าผิดพลาด ด้วยวิธีการซึ่งไม่ใช้ตรรกศาสตร์ คลุมเครือ (Non-fuzzy)	
	MAE \geq 0.5	MAE < 0.5	MAE \geq 0.5	MAE < 0.5
ระดับความพึงพอใจ 1	40.00	60.00	60.00	40.00
ระดับความพึงพอใจ 2	36.84	63.16	57.89	42.11
ระดับความพึงพอใจ 3	15.79	84.21	26.32	73.68
ระดับความพึงพอใจ 4	25.00	75.00	25.00	75.00
ระดับความพึงพอใจ 5	31.58	68.42	47.37	52.63
ภาพรวม	29.84	70.16	43.32	56.68

4.3 การอภิปรายผล

จากผลการประเมินความถูกต้องของส่วนอนุมานความรู้ สามารถสรุปสาเหตุที่ทำให้ค่ามาตรฐานในการประเมินอยู่ในเกณฑ์ต่ำ และเกิดความแตกต่างของระดับคะแนน ได้ดังนี้

4.3.1 ฐานความรู้ยังมีทรัพยากรไม่เพียงพอโดยเฉพาะคำศัพท์ ทั้งในส่วนของคลังคำที่มีความหมายเหมือนกันแต่เขียนต่างกัน คลังคำคุณศัพท์ และคลังคำกริยาวิเศษณ์ ซึ่งทั้งหมดนี้เป็นองค์ประกอบสำคัญที่ทำให้ระบบสามารถสกัดความรู้ได้ หากทรัพยากรเหล่านี้มีไม่เพียงพอ จะมีผลทำ

ให้ระบบไม่สามารถสกัดข้อมูลที่มีความเกี่ยวข้องได้ และส่งผลให้ค่า FN สูง จึงทำให้การคำนวณค่ามาตรฐานในการประเมินความถูกต้องอยู่ในเกณฑ์ค่อนข้างต่ำ และยังเป็นผลให้ระดับคะแนนที่ระบบสกัดได้มีค่าน้อยกว่าที่ผู้ใช้กำหนด

4.3.2 การใช้ทฤษฎีไวยากรณ์ไม่พึงบริบท ในขั้นตอนการวิเคราะห์เชิงโครงสร้างนั้น ผู้วิจัยยังกำหนดกลุ่มของกฎไวยากรณ์ได้ไม่ครอบคลุมทุกรูปแบบประโยคคำวิจารณ์ เนื่องจากรูปประโยคคำวิจารณ์ของผู้ใช้มีเป็นจำนวนมาก และมีความหลากหลายตามลักษณะการใช้ภาษาของผู้ใช้แต่ละคน ทำให้ระบบไม่สามารถสกัดข้อมูลที่มีความเกี่ยวข้องได้ ส่งผลให้ค่า FN สูง จึงทำให้การคำนวณค่ามาตรฐานในการประเมินความถูกต้องอยู่ในเกณฑ์ค่อนข้างต่ำ และยังเป็นผลให้ระดับคะแนนที่ระบบสกัดได้มีค่าน้อยกว่าที่ผู้ใช้กำหนด ดังจะเห็นได้อย่างชัดเจนในผลการคำนวณค่ามาตรฐานของการประเมินความถูกต้องในระดับความพึงพอใจที่ 1 ซึ่งพบว่ามีค่าอยู่ในเกณฑ์ต่ำกว่าระดับความพึงพอใจอื่น ๆ

4.3.3 การใช้ภาษาที่ไม่ถูกต้องตามหลักไวยากรณ์ภาษาอังกฤษ ทำให้ประโยคคำวิจารณ์ที่ควรนำมาสกัดเป็นความรู้ต้องถูกตัดออกไป จึงทำให้ความรู้หรือระดับคะแนนที่ระบบสกัดได้มีค่าน้อยกว่าความเป็นจริง

4.3.4 การกำหนดระดับคะแนนให้กับคำศัพท์ในฐานความรู้ของระบบ เป็นระดับคะแนนที่ได้จากการยืนยันของผู้เชี่ยวชาญด้านภาษา ซึ่งในความเป็นจริงผู้ใช้งานแต่ละคนอาจใช้คำศัพท์ในการแสดงความคิดเห็นที่ไม่สอดคล้องกับระดับคะแนนที่ได้จากการยืนยันของผู้เชี่ยวชาญด้านภาษา จึงทำให้ระดับคะแนนที่ระบบสกัดได้มีความแตกต่างจากที่ผู้ใช้กำหนด



บทที่ 5

สรุปและข้อเสนอแนะ

ในบทนี้กล่าวถึง สรุปผลการวิจัย ข้อจำกัดของการวิจัย การประยุกต์ผลการวิจัย และ ข้อเสนอแนะในการวิจัยครั้งต่อไป โดยมีรายละเอียด ดังนี้

5.1 สรุปผลการวิจัย

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อศึกษาและวิเคราะห์ขั้นตอนวิธีที่เกี่ยวข้องกับการสกัดความรู้แบบอัตโนมัติ แล้วนำมาออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย โดยใช้ข้อมูลด้านการท่องเที่ยว อันได้แก่ แหล่งที่พัก ประเภทที่พักเป็นกรณีศึกษา ในการพัฒนาระบบได้มุ่งเน้นการพัฒนาขั้นตอนวิธีในการวิเคราะห์ความหมายจากบทวิจารณ์ของนักท่องเที่ยวคนอื่น ๆ ที่ได้ไปแหล่งท่องเที่ยวมาแล้วมาสกัดเป็นความรู้ โดยนำเอาเทคนิคต่าง ๆ ที่มีอยู่ในปัจจุบัน เช่น ออนโทโลยีมาตรฐานด้านการท่องเที่ยวของเสียร์เพล์ พรานท์เนอร์ และบาชเลซเนอร์ และตัวแปลภาษา มาผสมผสานกับเทคนิคที่ได้คิดค้นขึ้นมาใหม่ อาทิ การคำนวณระดับคะแนนความพึงพอใจ ซึ่งเป็นความรู้ที่สกัดได้จากระบบ

จากการศึกษาขั้นตอนวิธีในการนำออนโทโลยีมาประยุกต์ใช้ในการสกัดความรู้ที่มีอยู่ในปัจจุบัน ผลการศึกษาทำให้ได้องค์ประกอบหลักของระบบที่จะพัฒนา ได้แก่ ตัววิเคราะห์คำศัพท์ ตัววิเคราะห์เชิงโครงสร้าง และตัววิเคราะห์เชิงความหมาย นอกจากนี้ การศึกษาปัจจัยที่มีผลต่อความพึงพอใจของนักท่องเที่ยว ยังทำให้ได้ส่วนการแสดงผลความรู้ของระบบที่ตอบสนองความต้องการของผู้ใช้มากที่สุด

ในการพัฒนาระบบ ผู้วิจัยทำการพัฒนาระบบดังกล่าวผ่านทางเว็บไซต์ โดยใช้ภาษา PHP ในการสร้างระบบ ใช้ชุดโปรแกรมเพาเวอร์ชาร์ต 3 (PowerCharts Version 3) เพื่อให้แสดงผลแผนภาพโครงสร้างต้นไม้ได้อย่างรวดเร็ว สวยงาม และเข้าใจง่าย ใช้มายเอสคิวแอล (MySQL) เป็นฐานข้อมูล ใช้โปรแกรมโปรเทจ รุ่น 3.4.1 ในการพัฒนาและปรับปรุงออนโทโลยี ซึ่งเป็นฐานความรู้ของระบบ และใช้อาร์เอพี (RDF API for PHP: RAP) ในการเชื่อมต่อระหว่างระบบกับออนโทโลยีที่พัฒนาขึ้น

ในส่วนของการประเมินความถูกต้องของการสกัดความรู้ นั้น ผู้วิจัยได้ทำการตรวจสอบความรู้ที่สกัดได้จากระบบใน 2 ประเด็น ได้แก่ (1) ความใกล้เคียงของระดับคะแนนที่ระบบสกัดได้กับระดับคะแนนโดยรวมทั้งผู้ใช้กำหนดและค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (2) นำผลการวิเคราะห์ประโยคคำวิจารณ์ที่ได้มาเป็นเกณฑ์ในการวัดความถูกต้องเทียบกับขั้นตอนวิธีในการสกัดความรู้แบบอัตโนมัติที่ได้พัฒนาขึ้นมา โดยการวัดความถูกต้องนั้นใช้วิธีวัดค่าที่เป็นมาตรฐานในการประเมิน ซึ่งในที่นี้จะมี 3 ค่า ได้แก่ ค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึก

สามารถสรุปผลการวิจัยได้ดังนี้

5.1.1 ในประเมินความถูกต้องของการสกัดความรู้ของระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย ผู้วิจัยได้เก็บข้อมูลบทวิจารณ์จำนวน 100 บทวิจารณ์ จากที่พักจำนวน 5 ที่พักซึ่งอยู่ในระดับความพึงพอใจที่ต่างกัน 5 ระดับ จำนวนระดับละ 20 บทวิจารณ์ จากเว็บไซต์เพื่อการท่องเที่ยว TripAdvisor.com โดยบทวิจารณ์เหล่านี้จำแนกได้เป็นข้อมูลที่สามารถสกัดได้และมีความเกี่ยวข้อง (TP) มีจำนวน 154 คำวิจารณ์; ไม่มีข้อมูลที่สามารถสกัดได้แต่ไม่มีความเกี่ยวข้อง (FP); ข้อมูลที่ไม่สามารถสกัดได้และไม่มีความเกี่ยวข้อง (TN) มีจำนวน 313 คำวิจารณ์ และข้อมูลที่ไม่สามารถสกัดได้แต่มีความเกี่ยวข้อง (FN) มีจำนวน 51 คำวิจารณ์ จากค่าของผลการทดสอบดังกล่าวสามารถคำนวณค่าความถูกต้องได้เท่ากับร้อยละ 90.15 ค่าความแม่นยำได้เท่ากับร้อยละ 100 ค่าความระลึกละเท่ากับร้อยละ 75.12 ค่าความแตกต่างของระดับคะแนนที่ระบบสกัดได้กับระดับคะแนนโดยรวมที่ผู้ใช้กำหนดเท่ากับ 0.37 คะแนนและค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองมีค่าเท่ากับ 0.466

5.1.2 ประเด็นปัญหาในการวิจัย สามารถสรุปได้ดังต่อไปนี้

1) ฐานความรู้ยังมีทรัพยากรไม่เพียงพอโดยเฉพาะคำศัพท์ ทั้งในส่วนของคลังคำที่มีความหมายเหมือนกันแต่เขียนต่างกัน คำคุณศัพท์ และคำกริยาวิเศษณ์ ซึ่งทั้งหมดนี้เป็นองค์ประกอบสำคัญที่ทำให้ระบบสามารถสกัดความรู้ได้

2) กลุ่มของกฎไวยากรณ์ในทฤษฎีไวยากรณ์ไม่พึงบริบทที่นำมาใช้ในขั้นตอนการวิเคราะห์เชิงโครงสร้าง ยังไม่ครอบคลุมทุกรูปแบบประโยคคำวิจารณ์ เนื่องจากรูปประโยคคำวิจารณ์ของผู้ใช้มีเป็นจำนวนมากและมีความหลากหลาย

ทั้งนี้ หากทรัพยากรในฐานความรู้ไม่เพียงพอ หรือกลุ่มของกฎไวยากรณ์ในทฤษฎีไวยากรณ์ไม่พึงบริบทยังไม่ครอบคลุมจะมีผลทำให้ระบบไม่สามารถสกัดข้อมูลที่มีความเกี่ยวข้องได้ ส่งผลให้การคำนวณค่ามาตรฐานในการประเมินความถูกต้องอยู่ในเกณฑ์ค่อนข้างต่ำ และยังทำให้ระดับคะแนนที่ระบบสกัดได้มีค่าน้อยกว่าที่ผู้ใช้กำหนดด้วย

3) การใช้ภาษาที่ไม่ถูกต้องตามหลักไวยากรณ์ภาษาอังกฤษ ทำให้ประโยคคำวิจารณ์ที่ควรนำมาสกัดเป็นความรู้ต้องถูกตัดออกไป จึงทำให้ความรู้หรือระดับคะแนนที่ระบบสกัดได้มีค่าน้อยกว่าความเป็นจริง

4) ระดับคะแนนของคำศัพท์ในฐานความรู้ของระบบ เป็นระดับคะแนนที่ได้จากการยืนยันของผู้เชี่ยวชาญด้านภาษา ซึ่งในความเป็นจริงผู้ใช้งานแต่ละคนอาจใช้คำศัพท์ในการแสดงความคิดเห็นที่ไม่สอดคล้องกับระดับคะแนนที่ได้จากการยืนยันของผู้เชี่ยวชาญด้านภาษา จึงทำให้ระดับคะแนนที่ระบบสกัดได้มีความแตกต่างจากที่ผู้ใช้กำหนด

5.2 ข้อจำกัดของการวิจัย

ในการพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทย มีข้อจำกัดในการวิจัย ดังนี้

5.2.1 ในการสกัดความรู้ของระบบ ผู้ใช้ต้องเขียนบทวิจารณ์เกี่ยวกับที่พักเป็นภาษาอังกฤษเท่านั้น หากเขียนบทวิจารณ์ในภาษาอื่น ระบบจะไม่สามารถประมวลผลได้

5.2.2 งานวิจัยนี้เป็นการสกัดความรู้จากบทวิจารณ์ที่เขียนขึ้นในภาษาธรรมชาติ ซึ่งเป็นภาษาที่มนุษย์ใช้ในการสื่อสาร โดยในความเป็นจริง ภาษาที่ใช้ในการสื่อสารนี้อาจไม่มีความถูกต้องตามหลักไวยากรณ์เสมอไป อีกทั้งยังขึ้นอยู่กับลักษณะการใช้ภาษาของแต่ละคน ดังนั้น กฎไวยากรณ์ไม่พึงบริบทที่ใช้ในงานวิจัยนี้จึงไม่อาจวิเคราะห์ประโยคได้ครอบคลุมทุกรูปแบบประโยคคำวิจารณ์

5.3 การประยุกต์ผลการวิจัย

ในการออกแบบและพัฒนาระบบสกัดและค้นคืนความรู้ผ่านระบบสารสนเทศทางภูมิศาสตร์ออนไลน์สำหรับธุรกิจท่องเที่ยวไทยนี้ ผู้ประกอบการธุรกิจอิเล็กทรอนิกส์สามารถนำผลการวิจัยไปประยุกต์ใช้เป็นต้นแบบการพัฒนาระบบในธุรกิจประเภทอื่น เพื่อเก็บและวิเคราะห์ข้อมูลเกี่ยวกับลูกค้าที่เป็นประโยชน์ต่อธุรกิจ แล้วนำข้อมูลเหล่านั้นมาวิเคราะห์และพัฒนาสินค้าหรือบริการ ตลอดจนวางนโยบายในการจัดการให้ตอบสนองความต้องการของลูกค้า เพื่อเปลี่ยนจากผู้บริโภคสู่การเป็นลูกค้าตลอดไปได้ในที่สุด

5.4 ข้อเสนอแนะในการวิจัยครั้งต่อไป

ในการสกัดความรู้แบบอัตโนมัตินั้น จำเป็นต้องพัฒนาฐานความรู้ให้มีความสมบูรณ์และเป็นปัจจุบันอยู่เสมอ เพื่อนำมาประกอบการพัฒนาระบบให้ผลลัพธ์มีความถูกต้องมากที่สุด ได้แก่ การเพิ่มเติมคำศัพท์ในฐานความรู้ให้มากขึ้น ทั้งในส่วนของคำที่มีความหมายเหมือนกันแต่เขียนต่างกัน คำคุณศัพท์ คำกริยาวิเศษณ์ รวมทั้งออนโทโลยีที่มีรายละเอียดมากขึ้น และกลุ่มของกฎไวยากรณ์ที่ใช้ในทฤษฎีไวยากรณ์ไม่พึงบริบทที่สามารถวิเคราะห์สำนวนหรือประโยคกำกวมได้มากขึ้น เนื่องจากข้อมูลเหล่านี้ล้วนเป็นปัจจัยที่ส่งผลต่อการคำนวณระดับคะแนนและส่งผลต่อการตัดสินใจของนักท่องเที่ยวในอนาคต จึงควรนำมาประกอบการพัฒนาระบบสกัดความรู้แบบอัตโนมัติทั้งสิ้น

อย่างไรก็ตาม หากระบบมีข้อมูลที่จำเป็นในการสกัดความรู้ทั้งหมดแล้ว สิ่งที่ควรพัฒนาถัดไปคือ การพัฒนาระบบให้มีการปรับปรุงฐานความรู้อัตโนมัติเมื่อพบคำศัพท์ใหม่ ๆ ในบทวิจารณ์ และการนำขั้นตอนวิธีที่พัฒนาให้มีความถูกต้องมากขึ้นมาประยุกต์ใช้ โดยการนำบทวิจารณ์จากข้อมูลหลายแหล่งมาวิเคราะห์และทดสอบ ตลอดจนพัฒนาระบบให้สามารถรองรับบทวิจารณ์ที่เขียนเป็นภาษาไทยได้ ซึ่งเป็นงานวิจัยที่จะดำเนินการต่อไปในอนาคต

บรรณานุกรม



บรรณานุกรม

- กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร. (2550). **พาณิชย์อิเล็กทรอนิกส์ คำขายจับไว เชื่อมโยงทั่วโลก ไทย ทั่วโลก** [ออนไลน์]. ได้จาก:
<http://www.mict.go.th/main.php?filename=ecom>
- กองประมวลผล และกองระบบสารสนเทศ. (2544). **การบริหารลูกค้าสัมพันธ์ (Customer Relationship Management: CRM)** [ออนไลน์]. ได้จาก:
<http://catadmin.cattelcom.com/training/crm.html>
- งานเทคโนโลยีประมวลผลข้อความ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. (2555). **เกี่ยวกับเล็กซิตรอน** [ออนไลน์]. ได้จาก: http://lexitron.nectec.or.th/2009_1/index.php?q=common_manager/aboutlex#history
- ณัฐพัชร์ ล้อประดิษฐ์พงษ์. (2549). **คู่มือสำรวจความพึงพอใจลูกค้า: Customer satisfaction survey handbook**. พิมพ์ครั้งที่ 2. กรุงเทพฯ: ประชุมทอง พรินต์ติ้ง กรุ๊ป.
- บุญเจริญ ศิริเนาวกุล. (2551). **ปัญญาประดิษฐ์: Artificial intelligence**. กรุงเทพฯ: ท้อป.
- มหาวิทยาลัยเกษตรศาสตร์. (2544). **การประมวลผลภาษาธรรมชาติ** [ออนไลน์]. ได้จาก:
http://www.ku.ac.th/magazine_online/nlp.html
- ยีน ภู่วรรณ และ ชัยยงค์ วงศ์ชัยสุวัฒน์. (2535). **การประมวลผลภาษาธรรมชาติ**. กรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าธนบุรี.
- สมชาย ปราการเจริญ. (2548). **ออนโทโลยี: ทางเลือกของการพัฒนาฐานความรู้ในรูปแบบเชิงเนื้อหา**. งานประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 1.
- ศิริวรรณ เสรีรัตน์ และคณะ. (2539). **คุณภาพในงานบริการ**. กรุงเทพฯ: ประชาชน.
- Alani, H., et al. (2003). **Automatic Ontology-based Knowledge Extraction from Web Documents**. IEEE Intelligent Systems. 18(1): 14-21.
- Bartolini, R., et al. (2006). **Multimedia Information Extraction in Ontology-based Semantic Annotation of Product Catalogues**. In Proceeding of Semantic Web Applications and Perspectives 2006. Pisa, Italy: CEUR-WS.
- Benamara, F., Cesarano, C., Picariello, A., Subrahmanian, V. S., Reforgiato, D. (2006). **Sentiment Analysis: Adjectives and Adverbs are Better than Adjectives Alone**. In Proceeding of the International Conference on Weblogs and Social Media (pp. 1-7).
- Chang, K. (2007). **Introduction to Geographic Information System**. Edition 4. McGraw Hill.

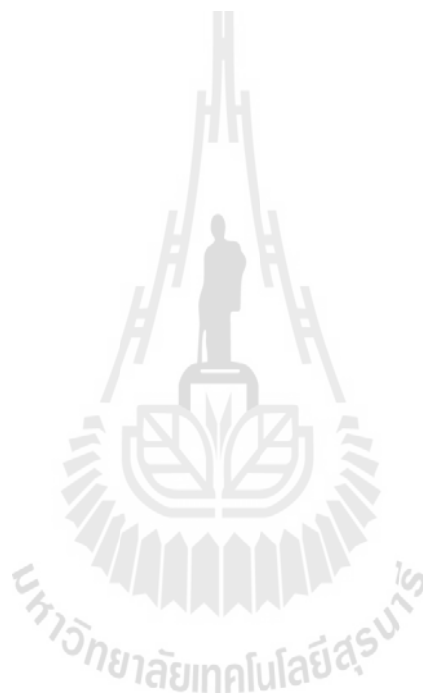
- Chanlekha, H., and Kawtrakul, A. (2001). **Thai Named Entity Extraction by incorporating Maximum Entropy Model with Simple Heuristic Information**. In Proceeding of International Joint Conference on Natural Language Processing (IJCNLP-04). Hainan, China: Springer.
- Chareonsuk, J., Sukvakree, T., and Kawtrakul, A. (2005). **Elementary Discourse unit Segmentation for Thai using Discourse Cue and Syntactic Information**. In Proceeding of the 9th National Computer Science and Engineering Conference (NCSEC 2005). Bangkok, Thailand: University of the Thai Chamber of Commerce.
- Chi, C. G-Q., and Qu, H. (2008). **Examining the structural relationships of destination image, tourist satisfaction and destination loyalty: An integrated approach**. *Tourism Management*. 29(4): 624-636.
- Cracolici, M. F., and Nijkamp, P. (2008). **The attractiveness and competitiveness of tourist destinations: A study of Southern Italian regions**. *Tourism Management*. 30(3): 336-344.
- Dave, K., Lawrence, S. and Pennock, D.M. (2003). **Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews**. In Proceeding of the 12th International Conference on World Wide Web (pp. 519-528). Budapest, Hungary: ACM.
- Davenport, T., and Prusak, L. (1998). **Working Knowledge: How Organisations Manage What They Know**. Boston: Harvard Business School Press.
- Del Bosque, I. R., and San Martín, H. (2008). **Tourism Research**. Great Britain: Elsevier. (*Annals of Tourism Research*. 35(2): 551-573)
- Fry, J., (2004). **Context-Free Grammars for English**. California, United States: San José State University.
- Gruber, T. R., (1993). **A Translation Approach to Portable Ontology Specifications**. *Knowledge Acquisition*. 5(2): 199-220.
- Haruechaiyasak, C., Srichaivattana, P., Kongyoung, S., and Damrongrat, C. (2004). **Automatic Thai Keyword Extraction from Categorized Text Corpus**. In The 1st ECTI Annual Conference. Pataya, Thailand: Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology association of Thailand.

- Horridge, M. (2011). **A Practical Guide to Building OWL Ontologies Using Prote'ge' 4 and CO-ODE Tools**. The University Of Manchester.
- Hepp, M., Siorpaes, K., and Bachlechner, D. (2006). **Towards the Semantic Web in e-Tourism: Lack of Semantics or Lack of Content?**. In Proceeding of the 3rd Annual European Semantic Web Conference (ESWC 2006). Budva, Montenegro: ESWC.
- Hsu, T-K., Tsai, Y-F., and Wu, H-H. (2009). **The preference analysis for tourist choice of destination: A case study of Taiwan**. *Tourism Management*. 30(2): 288–297.
- Hu, M. and Liu, B. (2004). **Mining and Summarizing Customer Reviews**. In Proceeding of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 168-177). Washington, USA: ACM.
- Imsoambut, A., Suktarachan, M., Yingsaree, W., and Kawtrakul A. (2005). **Country Report and Activity from Thailand: Ontology Construction and Maintenance System in Agricultural Domain**. In Proceeding of the International Joint Conference of the Asian Federation of Information Technology in Agriculture and the World Congress on Computer in Agriculture and Natural Resources (AFITA/WCCA 2004) (pp. 62-74). Bangkok, Thailand: Hydro and Agro Informatics Institute.
- Jang, J.-S. R., Sun, C.-T. and Mizutani, E. (1997). **Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence**. Upper Saddle River, NJ: Prentice Hall.
- Jurafsiy, D. and Martin, J. H. (2008). **Speech and Language Processing**. Edition 2. Printice Hall
- McGuinness, D. L., and Harmelen, F. V. (2004). **OWL Web Ontology Language Overview**. [On-line]. Available: <http://www.w3.org/TR/owl-features/>
- Miao, D., Duan, Q., Zhang, H., and Jiao, N. (2009). **Rough set based hybrid algorithm for text classification**. *Expert Systems with Applications*. 36(5): 9168–9174.
- Miniwatts, M. G. (2008). **World Internet Usage and Population Statistics** [On-line]. Available: <http://www.internetworldstats.com/stats3.htm>
- Nguyen, H. T., Walker, E. A. (2000). **A First Course in Fuzzy Logic**. Edition 2. Boca Raton, Chapman & Hall.

- Noroozi, A., and Fotouhi, Z. (2010). **The Influence of Semantic Web on Decision Making of Customers in Tourism Industry**. International Journal of Information Science and Management. (Special Issue I): 77-98.
- Noy, N. F., and McGuinness, L. (2003). **Ontology Evaluation: Not the Same as Schema Evolution**. Knowledge and Information Systems. [On-line]. Available: <http://www.springerlink.com/content/yjnvvn0vx5bcqnfj/>
- Peerawit, W., Yingsaeree, W., and Kawtrakul, A. (2004). **The Utilization of Closing Algorithm and Heuristic Information for Broken Character Segmentation**. In Proceeding of Cybernetics and Intelligent Systems (CIS2004). Singapore: IEEE.
- Pizam, A., Neumann, Y. and Reichel, A. (1978). **Dimensions of Tourist Satisfaction with a Destination area**. USA: Elsevier. (Annals of Tourism Research 5(3): 314–322)
- Popov, B., et al. (2003). **Towards Semantic Web Information Extraction**. In Proceeding of the 2nd International Semantic Web Conference (ISWC2003) Florida, USA: ISWC.
- Romero, F. P., Olivas, J. A., Genero, M., and Piattini, M. (2008). **Automatic extraction of the main terminology used in empirical software engineering through text mining techniques**. In Proceeding of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (pp. 357-358). Kaiserslautern, Germany: ACM.
- Roopa, J., Mladen, G., and Nalin, S. (2007). **Connecting Destinations with an Ontology-Based e-Tourism Planner**. In Proceeding of the 14th Annual Conference of IFITT, the International Federation for IT & Travel and Tourism (pp. 21-32). Ljubljana, Slovenia: Springer Vienna.
- Saaty, T. L. (2006). **Fundamentals of decision making and priority theory with the analytic hierarchy process**. Edition 2. Pittsburgh, PA: RWS.
- Salton, G. and Buckley, C. (1988). **Term Weighting Approaches in Automatic Text Retrieval**. Information Processing and Management. 24(5). pp 513-523.
- Siorpaes, K., Prantner, K., and Bachlechner, D. (2004). **Class Hierarchy for the e-Tourism Ontology Version 8** [PDF]. Galway, Ireland: Digital Enterprise Research Institute.

- Siricharoen, W.V. (2007). **E-commerce Adaptation Using Ontologies for E-tourism**. In Proceeding of 7th International Symposium on Communications and Information Technologies (pp. 1274-1279). Sydney, Australia: IEEE.
- Sirigayon, C., Chanlekha, H., and Kawtrakul, A. (2004). **Information Extraction for Agricultural Information Access**. In Proceeding of the International Joint Conference of the Asian Federation of Information Technology in Agriculture and the World Congress on Computer in Agriculture and Natural Resources (AFITA/WCCA 2004) (pp. 95-106). Bangkok, Thailand: Hydro and Agro Informatics Institute.
- Smitinand, T. (2001). **Thai Plant Names Dictionary**. The Forest Herbarium Royal Forest Department, Thailand.
- Sornlertlamvanich, V., Potipiti, T., and Charoenporn, T. (2000). **Automatic corpus-based Thai word extraction with the c4.5 learning algorithm**. In Proceeding of the 18th Conference on Computational Linguistics 2000 (pp. 802-807). Saarbrucken, Germany: Association for Computational Linguistics.
- Swartout, B., Patil, R., Knight, K., and Russ, T. (1996). **Toward Distributed Use of Large-Scale Ontologies**. Ontological Engineering. AAAI-97 Spring Symposium Series. 138-148.
- Taboada, M. and Grieve, J. (2004). **Analyzing Appraisal Automatically**. In Proceeding of the AAAI Spring Symposium. (pp. 158-161). Menlo Park, California: AAAI.
- Turney, P.D., (2002). **Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews**. In Proceeding of the 40th Annual Meeting of the Association for Computational Linguistics (ACL) (pp. 417-424). Philadelphia, America: Association for Computational Linguistics.
- Whitman, B. and Ellis, D. (2004). **Automatic record reviews**. In Proceeding of the 5th International Society for Music Information Retrieval (pp. 470-477). Barcelona, Spain: Universitat Pompeu Fabra.
- World Tourism Organization. (2007). **Yearbook of Tourism Statistics**. Madrid: WTO. (Yearbook of Tourism Statistics: Data 2001-2005 the 59th Edition)
- Wuttildercharoenwong, S. (2009) **Web based Expert Information Extraction**. Kasertsart University, Thailand.

- Yoon, Y., Uysal, M. (2005). **An Examination of the Effects of Motivation and Satisfaction on Destination Loyalty: A Structural Model**. *Tourism Management* 26 (pp. 45-56).
- Zheng, W. and Ye, Q. (2009). **Sentiment Classification of Chinese Traveler Reviews by Support Vector Machine Algorithm**. In *Proceeding of 2009 Third International Symposium on Intelligent Information Technology Application* (pp. 335-338). NanChang, China. IEEE.



ภาคผนวก ก

ตัวอย่างข้อมูลที่ได้จากระบบ

มหาวิทยาลัยเทคโนโลยีสุรนารี

1. ฐานความรู้

1.1 ออนโทโลยีด้านการท่องเที่ยว เป็นส่วนที่ใช้ในกระบวนการสกัดความรู้และจัดเก็บความรู้ทั้งหมดที่สกัดได้จากระบบ ประกอบด้วยโครงสร้างของคลาส (Class) และคุณสมบัติ (Properties) ของแต่ละคลาส ดังแสดงในตาราง ก.1-ก.15 ทั้งนี้ คุณสมบัติของคลาสแต่ละคลาส อาจประกอบไปด้วยคุณสมบัติทั้ง 2 ประเภท ได้แก่ คุณสมบัติที่มีค่าเป็นข้อมูล (Datatype properties) และคุณสมบัติที่มีค่าเป็นวัตถุ (Object properties)

ตารางที่ ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasContactData	Object properties	ContactData	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสข้อมูลการติดต่อ หมายถึง ข้อมูลที่ใช้ในการติดต่อของคลาสที่พัก อยู่ในคลาสข้อมูลการติดต่อ
hasPostalAddress	Object properties	PostalAddress	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสที่อยู่ หมายถึง ข้อมูลที่อยู่ของคลาสที่พัก อยู่ในคลาสที่อยู่
hasRoom	Object properties	Room	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสห้อง หมายถึง ห้องพักจะมีการเก็บรายละเอียดอยู่ในคลาสห้อง
hasGPSCoordinates	Object properties	GPSCoordinates	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสพิกัด หมายถึง ข้อมูลพิกัด และที่ตั้งของคลาสที่พัก อยู่ในคลาสพิกัด

ตารางที่ ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasOutdoorBathroom	Object properties	Bathroom	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสห้องน้ำ หมายถึง จะมีการเก็บรายละเอียดของห้องน้ำภายนอกห้องพักอยู่ในคลาสห้องน้ำ
hasRating	Object properties	Rating	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสการจัดอันดับ หมายถึง การจัดอันดับความพึงพอใจของคลาสที่พัก อยู่ในคลาสการจัดอันดับ
hasEvent	Object properties	Event	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสกิจกรรมพิเศษ หมายถึง การจัดอันดับความพึงพอใจของคลาสที่พัก อยู่ในคลาสกิจกรรมพิเศษ
hasDateTime	Object properties	DateTime	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสวัน เวลา เปิดทำการ หมายถึง การจัดอันดับความพึงพอใจของคลาสที่พัก อยู่ในคลาสวัน เวลา เปิดทำการ
hasActivity	Object properties	Activity	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสที่พักไปยังคลาสกิจกรรม หมายถึง การจัดอันดับความพึงพอใจของคลาสที่พัก อยู่ในคลาสกิจกรรม

ตารางที่ ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasCleaness	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงความสะอาดของที่พัก “มี” หรือ “ไม่มี”
hasDécor	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าการตกแต่งภายในที่พัก “มี” หรือ “ไม่มี”
hasSecurity	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าความปลอดภัย “มี” หรือ “ไม่มี”
hasValue	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าความคุ้มค่าของที่พัก “มี” หรือ “ไม่มี”
hasWellnessFacilities	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าสิ่งอำนวยความสะดวกต่าง ๆ “มี” หรือ “ไม่มี”
hasName	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายชื่อของที่พัก
spokenLanguage	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายภาษาที่ใช้ในการติดต่อสื่อสาร
hasAccommodationType	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายประเภทของที่พัก
hasHandicapAccessibility	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าคนพิการสามารถใช้บริการ “ได้” หรือ “ไม่ได้”

ตารางที่ ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
offersLaundryCleaning	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าบริการซักรีด “มี” หรือ “ไม่มี”
offersCot	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าบริการเตียงนอนสำหรับเด็ก “มี” หรือ “ไม่มี”
offersCrib	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าบริการเตียงนอนสำหรับเด็กอ่อน “มี” หรือ “ไม่มี”
offersShuttleService	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าบริการพาหนะขนส่งสาธารณะ “มี” หรือ “ไม่มี”
hasMoneyExchangeOffice	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าที่ให้บริการแลกเงิน “มี” หรือ “ไม่มี”
offersChildCare	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าบริการดูแลเด็ก “มี” หรือ “ไม่มี”
hasParking	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าจอดรถ “มี” หรือ “ไม่มี”
hasBeergarden	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าลานเบียร์ “มี” หรือ “ไม่มี”
hasElevator	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าลิฟต์ “มี” หรือ “ไม่มี”

ตารางที่ ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasPool	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าสระว่ายน้ำ “มี” หรือ “ไม่มี”
hasGarden	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าสวนสาธารณะ “มี” หรือ “ไม่มี”
hasEnglishNewspaper	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าหนังสือพิมพ์ภาษาอังกฤษ “มี” หรือ “ไม่มี”
PetsAllowed	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าอนุญาตให้นำสัตว์เลี้ยงมาได้ “มี” หรือ “ไม่มี”
hasFrontOffice	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าสำนักงานส่วนหน้า “มี” หรือ “ไม่มี”
hasRestaurant	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าร้านอาหาร “มี” หรือ “ไม่มี”
hasCafe	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าร้านขายเครื่องดื่ม “มี” หรือ “ไม่มี”
hasLocalFoodStall	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าร้านขายอาหารท้องถิ่น “มี” หรือ “ไม่มี”
hasSalon	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าร้านเสริมสวย “มี” หรือ “ไม่มี”

ตารางที่ ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasSolarium	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องกระจกเปิดรับแสงอาทิตย์ “มี” หรือ “ไม่มี”
hasBaggageRoom	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องเก็บสัมภาระ “มี” หรือ “ไม่มี”
hasSauna	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องซาวน่า “มี” หรือ “ไม่มี”
hasMessageStudio	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องบริการนวด “มี” หรือ “ไม่มี”
hasLobby	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องรับรอง (ล็อบบี้) “มี” หรือ “ไม่มี”
hasSpa	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องสปา “มี” หรือ “ไม่มี”
hasSteamBath	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าที่อาบน้ำแร่ “มี” หรือ “ไม่มี”
hasFitnessRoom	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องออกกำลังกาย “มี” หรือ “ไม่มี”
hasCocktail	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงคอกเทลเลาจ์ “มี” หรือ “ไม่มี”

ตารางที่ ก.1 คุณสมบัติทั้งหมดของคลาสที่พัก (Accommodations) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasBoard	Datatype properties	FullBoard, HalfBoard, Breakfast, AllInclusive	เป็นการระบุค่าคงที่ชนิดซิมโบล (Symbol) เพื่อใช้แสดงว่า FullBoard คือบริการอาหาร 3 มื้อ, HalfBoard คือบริการอาหาร 2 มื้อ (เช้า+กลางวันหรือเย็น), Breakfast คือบริการอาหารเช้า, AllInclusive คือบริการอาหารและเครื่องดื่ม
hasEmployees	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าพนักงาน “มี” หรือ “ไม่มี”
hasReception	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าพนักงานต้อนรับ “มี” หรือ “ไม่มี”
hasConcierge	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่ารองพนักงานเฝ้าประตู “มี” หรือ “ไม่มี”

ตารางที่ ก.2 คุณสมบัติทั้งหมดของคลาสห้องน้ำ (Bathroom)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasBellboy	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าพนักงานยกกระเป๋าในที่พัก “มี” หรือ “ไม่มี”
hasPhoto	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายภาพประกอบ
hasRating	Object properties	Rating	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรม ไปยังคลาสการจัดอันดับ หมายถึงการจัดอันดับความพึงพอใจของคลาสกิจกรรมอยู่ในคลาสการจัดอันดับ
hasHairDryer	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องเป่าผม “มี” หรือ “ไม่มี”
hasToilet	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าโถส้วม “มี” หรือ “ไม่มี”
hasShower	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าฝักบัวอาบน้ำ “มี” หรือ “ไม่มี”
hasWaterPressure	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าระบบปรับแรงน้ำ “มี” หรือ “ไม่มี”
hasSoapBar	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าที่สบู่ก้อน “มี” หรือ “ไม่มี”
hasRobes	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเสื้อคลุมอาบน้ำ “มี” หรือ “ไม่มี”

ตารางที่ ก.2 คุณสมบัติทั้งหมดของคลาสห้องน้ำ (Bathroom) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasBathtub	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าอ่างอาบน้ำ “มี” หรือ “ไม่มี”

ตารางที่ ก.3 คุณสมบัติทั้งหมดของคลาสห้อง (Room)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasRating	Object properties	Rating	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสห้องไปยังคลาสการจัดอันดับ หมายถึงการจัดอันดับความพึงพอใจของคลาสห้อง อยู่ในคลาส การจัดอันดับ
hasRoomType	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายประเภทห้องพัก
smokingAllowed	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าอนุญาตให้สูบบุหรี่ภายในห้องพัก “มี” หรือ “ไม่มี”
hasInternetAccess	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าอินเทอร์เน็ต “มี” หรือ “ไม่มี”
hasVCR	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องเล่นวีดีทัศน์ “มี” หรือ “ไม่มี”
hasHotwater	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าน้ำร้อน “มี” หรือ “ไม่มี”

ตารางที่ ก.3 คุณสมบัติทั้งหมดของคลาสห้อง (Room) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasTVSet	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องโทรทัศน์ “มี” หรือ “ไม่มี”
hasTelephone	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องโทรศัพท์ “มี” หรือ “ไม่มี”
hasAirCondition	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องปรับอากาศ “มี” หรือ “ไม่มี”
hasFaxMachine	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องแฟกซ์ “มี” หรือ “ไม่มี”
hasArea	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าบริเวณภายในห้อง “มี” หรือ “ไม่มี”

ตารางที่ ก.4 คุณสมบัติทั้งหมดของคลาสห้องพัก (Guestroom)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasIndoorBathroom	Object properties	Bathroom	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสห้องพัก ไปยังคลาสห้องน้ำ หมายถึงจะมีการเก็บรายละเอียดของห้องน้ำภายในห้องพักอยู่ในคลาสห้องพัก

ตารางที่ ก.4 คุณสมบัติทั้งหมดของคลาสห้องพัก (Guestroom) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasRating	Object properties	Rating	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสห้องพัก ไปยังคลาสการจัดอันดับ หมายถึงการจัดอันดับความพึงพอใจของคลาสห้องพัก อยู่ในคลาสการจัดอันดับ
hasBedType	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายประเภทเตียงนอน
hasIce	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าน้ำแข็ง “มี” หรือ “ไม่มี”
hasSpringWater	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าน้ำเปล่า “มี” หรือ “ไม่มี”
hasCableTV	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเคเบิลทีวี “มี” หรือ “ไม่มี”
hasMicrowave	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องเตาอบไมโครเวฟ “มี” หรือ “ไม่มี”
hasBalcony	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าระเบียงหน้าต่าง “มี” หรือ “ไม่มี”
hasSafe	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าตู้นิรภัย “มี” หรือ “ไม่มี”
hasRefrigerator	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าตู้เย็น “มี” หรือ “ไม่มี”

ตารางที่ ก.4 คุณสมบัติทั้งหมดของคลาสห้องพัก (Guestroom) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasBed	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเตียงนอน “มี” หรือ “ไม่มี”
hasView	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าห้องพักสามารถชมทิวทัศน์ได้ “มี” หรือ “ไม่มี”
hasAlarmClock	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่านาฬิกาปลุก “มี” หรือ “ไม่มี”
hasBedsheet	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าผ้าปูเตียง “มี” หรือ “ไม่มี”
hasMinibar	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่ามีมินิบาร์ “มี” หรือ “ไม่มี”
hasSlippers	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่ารองเท้าแตะใส่ในห้อง “มี” หรือ “ไม่มี”
hasTerrace	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าที่นั่งเล่นกลางแจ้ง “มี” หรือ “ไม่มี”
hasKitchenetteDrawer	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าลิ้นชักเครื่องครัว “มี” หรือ “ไม่มี”
hasTeaCoffeeEquipment	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าอุปกรณ์ชงชาและกาแฟ “มี” หรือ “ไม่มี”
hasBreakfast	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าอาหารเช้า “มี” หรือ “ไม่มี”

ตารางที่ ก.5 คุณสมบัติทั้งหมดของคลาสห้องประชุม (ConferenceRoom)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasRating	Object properties	Rating	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสห้องประชุม ไปยังคลาสการจัดอันดับ หมายถึงการจัดอันดับความพึงพอใจของคลาสห้อง อยู่ในคลาสการจัดอันดับ
hasVideoProjector	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเครื่องวิดีโอโพรเจกเตอร์ “มี” หรือ “ไม่มี”
hasScreen	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าจอภาพ “มี” หรือ “ไม่มี”
hasFlipChart	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าป้ายชื่อหน้าห้อง “มี” หรือ “ไม่มี”
hasLectern	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าโพลีเดียม “มี” หรือ “ไม่มี”
hasVideoConference System	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าระบบการประชุมทางไกลผ่านจอภาพ “มี” หรือ “ไม่มี”
litByNaturalDaylight	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าระบบส่องสว่างด้านแสงธรรมชาติ “มี” หรือ “ไม่มี”
hasStage	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าเวที “มี” หรือ “ไม่มี”

ตารางที่ ก.5 คุณสมบัติทั้งหมดของคลาสห้องประชุม (ConferenceRoom) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasAudioEquipment	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้แสดงว่าอุปกรณ์เกี่ยวกับเครื่องเสียง “มี” หรือ “ไม่มี”

ตารางที่ ก.6 คุณสมบัติทั้งหมดของคลาสกิจกรรม (Activity)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
canBeDoneAt	Object properties	Accommodation	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรม ไปยังคลาสที่พัก หมายถึงกิจกรรมภายในที่พักรสามารถทำได้ที่บริเวณใด
hasRating	Object properties	Rating	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรมไปยังคลาสการจัดอันดับ หมายถึงการจัดอันดับความพึงพอใจของคลาสกิจกรรม อยู่ในคลาสการจัดอันดับ
hasPhoto	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายภาพประกอบ
hasName	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายชื่อของกิจกรรม

ตารางที่ ก.7 คุณสมบัติทั้งหมดของคลาสสถานที่อำนวยความสะดวก (Amenity)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasActivity	Object properties	Activity	เป็น คุณสมบัติที่ เชื่อมโยง ความสัมพันธ์ระหว่างคลาส สถานที่อำนวยความสะดวก ไปยัง คลาสกิจกรรม หมายถึงกิจกรรมที่สามารถทำได้ ของคลาสสถานที่อำนวยความสะดวก อยู่ในคลาสิกิจกรรม
hasRating	Object properties	Rating	เป็น คุณสมบัติที่ เชื่อมโยง ความสัมพันธ์ระหว่างคลาส สถานที่อำนวยความสะดวก ไปยัง คลาสการจัดอันดับ หมายถึงการจัดอันดับความพึงพอใจของคลาสสถานที่อำนวยความสะดวก อยู่ในคลาสิกการจัดอันดับ
hasGPSCoordinates	Object properties	GPSCoordinates	เป็น คุณสมบัติที่ เชื่อมโยง ความสัมพันธ์ระหว่างคลาส สถานที่อำนวยความสะดวก ไปยัง คลาสพิกัด หมายถึงข้อมูลพิกัดและที่ตั้งของคลาสสถานที่อำนวยความสะดวก อยู่ในคลาสิกพิกัด
hasPostalAddress	Object properties	PostalAddress	เป็น คุณสมบัติที่ เชื่อมโยง ความสัมพันธ์ระหว่างคลาส สถานที่อำนวยความสะดวก ไปยัง คลาสที่อยู่ หมายถึงข้อมูลที่อยู่ของคลาสสถานที่อำนวยความสะดวก อยู่ในคลาสิกที่อยู่
hasContactData	Object properties	ContactData	เป็น คุณสมบัติที่ เชื่อมโยง ความสัมพันธ์ระหว่างคลาส สถานที่อำนวยความสะดวก ไปยัง คลาสข้อมูลการติดต่อ หมายถึงข้อมูลที่ใช้ในการติดต่อของคลาสสถานที่อำนวยความสะดวก อยู่ในคลาสิกข้อมูลการติดต่อ

ตารางที่ ก.7 คุณสมบัติทั้งหมดของคลาสสถานที่อำนวยความสะดวก (Amenity) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasDateTime	Object properties	DateTime	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่อำนวยความสะดวก ไปยังคลาสวัน เวลา เปิดทำการ หมายถึงข้อมูลวัน เวลา ทำการของสถานที่อำนวยความสะดวก อยู่ในคลาสวัน เวลา เปิดทำการ
hasEvent	Object properties	Event	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่อำนวยความสะดวก ไปยังคลาสกิจกรรมพิเศษ หมายถึงกิจกรรมพิเศษของสถานที่อำนวยความสะดวก อยู่ในคลาสกิจกรรมพิเศษ
hasPhoto	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายภาพประกอบ
hasName	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายชื่อของสถานที่อำนวยความสะดวก
hasType	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายประเภทของสถานที่อำนวยความสะดวก

ตารางที่ ก.8 คุณสมบัติทั้งหมดของคลาสสถานที่ท่องเที่ยว (Attraction)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasActivity	Object properties	Activity	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสกิจกรรม หมายถึงกิจกรรมที่สามารถทำได้ ของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสกิจกรรม
hasRating	Object properties	Rating	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสการจัดอันดับ หมายถึงการจัดอันดับความพึงพอใจของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสการจัดอันดับ
hasGPSCoordinates	Object properties	GPSCoordinates	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสพิกัด หมายถึงข้อมูลพิกัดและที่ตั้งของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสพิกัด
hasPostalAddress	Object properties	PostalAddress	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสที่อยู่ หมายถึงข้อมูลที่อยู่ของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสที่อยู่
hasContactData	Object properties	ContactData	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสข้อมูลการติดต่อ หมายถึงข้อมูลที่ใช้ในการติดต่อของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสข้อมูลการติดต่อ

ตารางที่ ก.8 คุณสมบัติทั้งหมดของคลาสสถานที่ท่องเที่ยว (Attraction) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasDateTime	Object properties	DateTime	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสวัน เวลา เปิดทำการ หมายถึงข้อมูลวัน เวลา ทำการ ของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสวัน เวลา เปิดทำการ
hasEvent	Object properties	Event	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสกิจกรรมพิเศษ หมายถึงกิจกรรมพิเศษของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสกิจกรรมพิเศษ
hasTicket	Object properties	Ticket	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสสถานที่ท่องเที่ยว ไปยังคลาสตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว หมายถึงข้อมูลตั๋วเข้าชมสถานที่ท่องเที่ยวของคลาสสถานที่ท่องเที่ยว อยู่ในคลาสตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว
hasPhoto	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายภาพประกอบ
hasName	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายชื่อของสถานที่ท่องเที่ยว

ตารางที่ ก.8 คุณสมบัติทั้งหมดของคลาสสถานที่ท่องเที่ยว (Attraction) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasType	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายประเภทของสถานที่ท่องเที่ยว

ตารางที่ ก.9 คุณสมบัติทั้งหมดของคลาสข้อมูลการติดต่อ (ContactData)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasEMail	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายอีเมล
hasFaxNumber	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายหมายเลขโทรสาร
hasTelephoneNumber	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายหมายเลขโทรศัพท์
hasWebsite	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายเว็บไซต์

ตารางที่ ก.10 คุณสมบัติทั้งหมดของคลาสวัน เวลา เปิดทำการ (DateTime)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasStartDate	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายวันเปิดทำการ
hasEndDate	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายวันปิดทำการ
hasEndTime	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายเวลาปิดทำการ
hasStartTime	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายเวลาเปิดทำการ

ตารางที่ ก.10 คุณสมบัติทั้งหมดของคลาสวัน เวลา เปิดทำการ (DateTime) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasWeekday	Datatype properties	Monday, Tuesday, Wednesday, Thursday, Friday	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายวันทำการที่เป็นวันธรรมดา (จันทร์ - ศุกร์)

ตารางที่ ก.11 คุณสมบัติทั้งหมดของคลาสกิจกรรมพิเศษ (Event)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasActivity	Object properties	Activity	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรมพิเศษ ไปยังคลาสกิจกรรม หมายถึงกิจกรรมที่สามารถทำได้ ของคลาสกิจกรรมพิเศษ อยู่ในคลาสกิจกรรม
hasGPSCoordinates	Object properties	GPSCoordinates	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรมพิเศษ ไปยังคลาสพิกัด หมายถึงข้อมูลพิกัดและที่ตั้งของคลาสกิจกรรมพิเศษ อยู่ในคลาสพิกัด
hasPostalAddress	Object properties	PostalAddress	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรมพิเศษ ไปยังคลาสที่อยู่ หมายถึงข้อมูลที่อยู่ของคลาสกิจกรรมพิเศษ อยู่ในคลาสที่อยู่
hasContactData	Object properties	ContactData	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรมพิเศษ ไปยังคลาสข้อมูลการติดต่อ หมายถึงข้อมูลที่ใช้ในการติดต่อของคลาสกิจกรรมพิเศษ อยู่ในคลาสข้อมูลการติดต่อ

ตารางที่ ก.11 คุณสมบัติทั้งหมดของคลาสกิจกรรมพิเศษ (Event) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasDateTime	Object properties	DateTime	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรมพิเศษ ไปยังคลาสวัน เวลา เปิดทำการ หมายถึงข้อมูล วัน เวลา ทำการ ของคลาสกิจกรรมพิเศษ อยู่ในคลาสวัน เวลา เปิดทำการ
hasTicket	Object properties	Ticket	เป็นคุณสมบัติที่เชื่อมโยงความสัมพันธ์ระหว่างคลาสกิจกรรมพิเศษ ไปยังคลาสตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว หมายถึงข้อมูลตั๋วเข้าชมสถานที่ท่องเที่ยวของคลาสกิจกรรมพิเศษ อยู่ในคลาสตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว
hasPhoto	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายภาพประกอบ
hasName	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายชื่อของกิจกรรมพิเศษ
hasType	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายประเภทของกิจกรรมพิเศษ

ตารางที่ ก.12 คุณสมบัติทั้งหมดของคลาสพิกัด (GPSCoordinates)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasLongitude	Datatype properties	float	เป็นการระบุค่าคงที่ชนิด ตัวเลขทศนิยม (Float) เพื่อใช้อธิบายเส้นลองจิจูด
hasLatitude	Datatype properties	float	เป็นการระบุค่าคงที่ชนิด ตัวเลขทศนิยม (Float) เพื่อใช้อธิบายเส้นละติจูด

ตารางที่ ก.13 แสดงคุณสมบัติทั้งหมดของคลาสที่อยู่ (PostalAddress)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasHouseNumber	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายบ้านเลขที่
hasStreet	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายถนน
hasTambon	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายตำบล
hasDistrict	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายอำเภอ
hasProvince	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายจังหวัด

ตารางที่ ก.14 คุณสมบัติทั้งหมดของคลาสการจัดอันดับ (Rating)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
Excellent	Datatype properties	Integer	เป็นการระบุค่าคงที่ชนิด ตัวเลขจำนวนเต็ม (Integer) เพื่อใช้อธิบายระดับความพึงพอใจในระดับดีมาก
Good	Datatype properties	Integer	เป็นการระบุค่าคงที่ชนิด ตัวเลขจำนวนเต็ม (Integer) เพื่อใช้อธิบายระดับความพึงพอใจในระดับดี

ตารางที่ ก.14 คุณสมบัติทั้งหมดของคลาสการจัดอันดับ (Rating) (ต่อ)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
Average	Datatype properties	Integer	เป็นการระบุค่าคงที่ชนิด ตัวเลขจำนวนเต็ม (Integer) เพื่อใช้อธิบายระดับความพึงพอใจในระดับปานกลาง
Bad	Datatype properties	Integer	เป็นการระบุค่าคงที่ชนิด ตัวเลขจำนวนเต็ม (Integer) เพื่อใช้อธิบายระดับความพึงพอใจในระดับแย่
Worse	Datatype properties	Integer	เป็นการระบุค่าคงที่ชนิด ตัวเลขจำนวนเต็ม (Integer) เพื่อใช้อธิบายระดับความพึงพอใจในระดับแย่มาก

ตารางที่ ก.15 คุณสมบัติทั้งหมดของคลาสตั๋วเข้าชมการแสดงหรือสถานที่ท่องเที่ยว (Ticket)

ชื่อคุณสมบัติ	ประเภท	ช่วงค่าที่อนุญาต	คำอธิบาย
hasName	Datatype properties	String	เป็นการระบุค่าคงที่ชนิดข้อความ (String) เพื่อใช้อธิบายชื่อของตั๋วการแสดงหรือที่ท่องเที่ยวต่าง ๆ
hasPrice	Datatype properties	Integer	เป็นการระบุค่าคงที่ชนิด ตัวเลขจำนวนเต็ม (Integer) เพื่อใช้อธิบายราคาของตั๋วการแสดงหรือที่ท่องเที่ยว

1.2 คลังคำเหมือน เป็นแหล่งรวบรวมคำต่าง ๆ ที่มีความหมายเหมือนกัน แต่เขียนต่างกัน รวมถึงการตรวจสอบคำผิดเพื่อเปลี่ยนให้เป็นคำที่ถูกต้อง ดังแสดงในตารางที่ ก.16 โดยจะมีการกำหนดคำที่เป็นตัวแทนของกลุ่มคำเหมือนแต่ละกลุ่ม เพื่อใช้ในการแปลงคำศัพท์ให้อยู่ในรูปแบบที่เหมาะสม สำหรับการสกัดความรู้

ตารางที่ ก.16 คลังคำเหมือนของระบบ

ตัวแทน	คำเหมือน
'	,
advertisement	ad
aircondition	ac, a.c., a/c, air con, air conditioning, airconditioning, air conditioned, airconditioned, aircon, air con, air conditioner, airconditioner, air conditions, airconditions
alarmclock	alarm clock, clock, clocks
and	&
are not	aint, ain't, ainot
audioequipment	audio equipment, stereo
bathroom	bath room
bathtub	basin, bath tub, bath, sink, private bath
bed	beds, mattress, mattresses, matress
bedbug	bed bugs, bedbugs
bedsheet	bed sheet, bedsheets, bed sheets, bed cover, bedcover, bed spread, tick, ticks, pillow, quilt, covering, bed linen, bedlinen, sheets
bellboy	bell boy, bellboys, bellguy, bellhop, bellmen
breakfast	continental breakfast, english breakfast, brkfst
cabletv	satellite tv, cable tv
calm	calmed, calming, calms
can not	cant, can't, cannot, cannt
childcare	child, childs, children
clean	cleaned, cleans
cleanness	cleanliness, cleanness

ตารางที่ ก.16 คลังคำเหมือนของระบบ (ต่อ)

ตัวแทน	คำเหมือน
clear	cleared, clearing, clears
could not	couldnt, couldn't, couldnot
decor	decorated, decorating, decoration, decorations, decorative
delicious	deliciously, deliciousness
did not	didnt, didn't, didnot, did'nt
do not	dont, don't, donot
does not	doesnt, doesn't, doesnot
e	é
employee	employees, staff, wait staff
englishnewspaper	english newspaper
fair	fairly, fairs
faxmachine	fax machine
first	1st
fitness	fitness room, gym, exercise room
flipchart	flip chart, label
frontoffice	front office, front desk, clerk
frustrate	frustrating, frustrated, frustrates
functional	functioned, functioning, functionally
hairdryer	hair dryer
handicap	cripple, handicapped, defective
have	've
hotwater	hot water, water heater
i am	Im, Im
internetaccess	wireless internet, internet access, wifi, wirelessconnection, wireless lan, access point, internet, free internet
kitchenettedrawer	kitchenette drawer
laudrycleaning	laudry cleaning, laundry cleaning, laundry
lobby	lounge, foyer, executive lobby, executive lounge
localfoodstall	food shop, local food stall, local shops

ตารางที่ ก.16 คลังคำเหมือนของระบบ (ต่อ)

ตัวแทน	คำเหมือน
location	located
love	loved, loving, loves
massagestudio	massage
moneyexchangeoffice	money exchange
not	n't
outoforder	out of order
overprice	overprices, overpriced, overpricing
parking	car park, parking place, parking lot
petsallowed	pet, pets, pets allowed, pets allow
pool	pond
reception	reception, reception staff, receptionist, receptionists, owner, owners, proprietor, host, service
refrigerator	fridge, freezer, cooler
restaurant	hotel restaurant, eatery, resturants, restaurants, ethnic restaurants
ripoff	rip off
robes	bathrobe, bath robe, bathing gown, bathing wrap, robe
room	rooms
roomtype	rooms, room, suite, single room, basement room, bedroom, bed room
salon	beauty shop, beauty parlor, hairdressing salon, beauty salon
security	secure, secured
service	services
shuttleservice	shuttle service
slippers	slipper
smell	smells, smelled, smelling, smelly
smokingallowed	smoking allowed
soapbar	soap bar

ตารางที่ ก.16 คลังคำเหมือนของระบบ (ต่อ)

ตัวแทน	คำเหมือน
spokenlanguage	spoken language, speaking language, language speaking, spoke language, speaks language
springwater	spring water
stain	stains
stayaway	stay away
steambath	steam bath
teacoffeeequipment	tea coffee equipment, coffee equipment, tea, coffee, teacoffee facilites, tea facilites, coffee facilites, tea/coffee facilites
telephone	phone
thereare	there are
tvset	t.v., television, tv
unwelcome	unwelcoming
value	rate, prices, price, values
videoconferencesystem	video conference system, vdo conference, video conference
videoprojector	video projector, vdo projector, projector
view	viewed, viewing, views
was not	wasnt, wasn't, wasnot
waterpressure	pressure, water pressure
welcome	welcoming
welldone	well done
wellnessfacilities	facilities
wellpresent	well presented, well present
will not	wont, won't, willnot
wornout	worn out
would not	wouldnt, wouldn't, wouldnot

1.3 คลังคำศัพท์ เป็นแหล่งรวบรวมคำศัพท์พร้อมทั้งระดับคะแนนที่จำเป็นในการสกัดความรู้ ซึ่งจัดเก็บอยู่ในรูปแบบภาษาเอ็กซ์เอ็มแอล (XML) ที่เป็นภาษาในการอธิบายข้อมูล โดยมีการใช้ป้ายระบุ (Tag) เพื่อบอกหน้าที่และประเภทข้อมูลของส่วนต่าง ๆ ในเอกสารนั้น ทำให้การประมวลผลเอกสารเป็นไปโดยง่ายและไม่จำเป็นต้องอาศัยมนุษย์เพื่อตีความเอกสาร ทั้งนี้ ผู้วิจัยได้อธิบายป้ายระบุที่ใช้ในคลังคำศัพท์ไว้ดังตารางที่ ก.17 และแสดงคลังคำศัพท์ของระบบในภาพที่ ก.1

ตารางที่ ก.17 คำอธิบายป้ายระบุที่ใช้ในคลังคำศัพท์

ชื่อป้ายระบุ	คำอธิบาย
<words>	ใช้อธิบายกลุ่มคำศัพท์ทั้งหมดที่อยู่ในคลังคำศัพท์ของระบบ
<basics>	ใช้อธิบายกลุ่มคำคุณศัพท์ทั้งหมด
<basic>	ใช้อธิบายคำคุณศัพท์ที่มีการกำหนดระดับคะแนนไว้
<vocab>	ใช้อธิบายคำคุณศัพท์
<rate>	ใช้อธิบายระดับคะแนน
<advances>	ใช้อธิบายกลุ่มคำกริยาวิเศษณ์ทั้งหมด
<advance>	ใช้อธิบายคำกริยาวิเศษณ์ที่กำหนดการเพิ่ม-ลดระดับคะแนนไว้
<vocab>	ใช้อธิบายคำกริยาวิเศษณ์
<rateless>	ใช้อธิบายการเพิ่ม-ลดระดับคะแนน ในกรณีที่ระดับคะแนนรวมมีค่าน้อยกว่า 3
<ratemoreequal>	ใช้อธิบายการเพิ่ม-ลดระดับคะแนน ในกรณีที่ระดับคะแนนรวมมีค่ามากกว่าหรือเท่ากับ 3
<specials>	ใช้อธิบายกลุ่มคำกริยาวิเศษณ์แบบพิเศษทั้งหมด
<special>	ใช้อธิบายคำกริยาวิเศษณ์แบบพิเศษที่กำหนดการเพิ่ม-ลดระดับคะแนนไว้
<vocab>	ใช้อธิบายคำกริยาวิเศษณ์แบบพิเศษ
<rateless>	ใช้อธิบายการเพิ่ม-ลดระดับคะแนน ในกรณีที่ระดับคะแนนรวมมีค่าน้อยกว่า 3
<ratemoreequal>	ใช้อธิบายการเพิ่ม-ลดระดับคะแนน ในกรณีที่ระดับคะแนนรวมมีค่ามากกว่าหรือเท่ากับ 3

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<words>
<basics>
  <basic>
    <vocab>dark</vocab>
    <rate>2</rate>
  </basic>
  <basic>
    <vocab>dirty</vocab>
    <rate>2</rate>
  </basic>
  <basic>
    <vocab>grubby</vocab>
    <rate>2</rate>
  </basic>
  <basic>
    <vocab>worse</vocab>
    <rate>1</rate>
  </basic>
  <basic>
    <vocab>acceptable</vocab>
    <rate>3</rate>
  </basic>
  <basic>
    <vocab>unacceptable</vocab>
    <rate>1</rate>
  </basic>
  <basic>
    <vocab>subpar</vocab>
    <rate>2</rate>
  </basic>
  <basic>
    <vocab>sparse</vocab>
    <rate>2</rate>
  </basic>
  <basic>
    <vocab>moderate</vocab>
    <rate>3</rate>
  </basic>
  <basic>
    <vocab>excellent</vocab>
    <rate>5</rate>
  </basic>
  <basic>
    <vocab>good</vocab>
    <rate>4</rate>
  </basic>
  <basic>
    <vocab>minimal</vocab>
    <rate>1</rate>
  </basic>
  <basic>
    <vocab>minimum</vocab>
    <rate>1</rate>
  </basic>
</basics>
</words>
```

```
<vocab>bad</vocab>
<rate>2</rate>
</basic>
<basic>
  <vocab>badly</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>far</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>smooth</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>friendly</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>unfriendly</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>helpful</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>weird</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>passable</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>adequate</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>inadequate</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>sympathetic</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>fair</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>comfortable</vocab>
  <rate>4</rate>
</basic>
<basic>
```

```
<vocab>fine</vocab>
<rate>4</rate>
</basic>
<basic>
  <vocab>has</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>had</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>incredible</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>ideal</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>servicemind</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>serviceminded</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>have</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>thereare</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>no</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>average</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>awful</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>basic</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>best</vocab>
  <rate>5</rate>
</basic>
<basic>
```

```
<vocab>better</vocab>
<rate>4</rate>
</basic>
<basic>
  <vocab>clean</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>unclean</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>comfort</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>contented</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>cool</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>cold</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>dated</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>decent</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>difficultcommunicate</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>disappoint</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>disappointing</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>dreadful</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>efficient</vocab>
  <rate>5</rate>
</basic>
<basic>
```



```
<vocab>expensive</vocab>
<rate>2</rate>
</basic>
<basic>
  <vocab>fab</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>fantastic</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>fresh</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>great</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>grotty</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>happy</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>cheerful</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>happily</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>hate</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>impolite</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>polite</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>inconsiderate</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>inconvenience</vocab>
  <rate>2</rate>
</basic>
<basic>
```

```
<vocab>convenience</vocab>
<rate>4</rate>
</basic>
<basic>
  <vocab>convenient</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>neat</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>suitable</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>infestations</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>infest</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>kind</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>loud</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>love</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>lovely</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>modern</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>stylish</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>prime</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>narrow</vocab>
  <rate>2</rate>
</basic>
<basic>
```

```
<vocab>nice</vocab>
<rate>4</rate>
</basic>
<basic>
  <vocab>noisy</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>ok</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>old</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>overprice</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>perfect</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>pleasant</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>unpleasant</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>unpleasing</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>unpleasantly</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>pleased</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>quiet</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>feeble</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>prompt</vocab>
  <rate>4</rate>
</basic>
<basic>
```

```
<vocab>reasonable</vocab>
<rate>3</rate>
</basic>
<basic>
  <vocab>civil</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>rubbish</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>rude</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>tatty</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>rusty</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>shame</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>shiny</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>smart</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>smell</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>smelt</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>smooth</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>spectacular</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>stinking</vocab>
  <rate>1</rate>
</basic>
<basic>
```

```
<vocab>stink</vocab>
<rate>1</rate>
</basic>
<basic>
  <vocab>stunk</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>sullen</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>surprise</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>terrible</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>unhelpful</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>welcoming</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>welcome</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>unwelcoming</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>unwelcome</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>worst</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>frustrate</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>wornout</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>cockroach</vocab>
  <rate>1</rate>
</basic>
<basic>
```

```
<vocab>roach</vocab>
<rate>1</rate>
</basic>

<basic>
  <vocab>bedbug</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>misleading</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>well</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>nasty</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>stained</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>stain</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>broken</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>cracked</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>unprofessional</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>disgusting</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>repulsive</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>laughable</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>stupid</vocab>
  <rate>1</rate>
</basic>
```

```
<basic>
  <vocab>dingy</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>shabby</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>desperate</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>plain</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>abusive</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>lousy</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>creepy</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>normal</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>cheap</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>okay</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>stuffy</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>tiny</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>presentable</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>kindness</vocab>
  <rate>4</rate>
</basic>
```

```
<basic>
  <vocab>plentiful</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>fungus</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>molds</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>worth</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>varied</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>gorgeous</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>ornate</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>beautiful</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>fluent</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>superb</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>functional</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>complimentary</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>discreet</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>fabulous</vocab>
  <rate>5</rate>
</basic>
```



```
<basic>
  <vocab>regretful</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>regret</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>scented</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>calm</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>fashionable</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>chic</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>chicness</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>graceful</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>professional</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>fit</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>relaxing</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>relax</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>phenomenal</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>impeccable</vocab>
  <rate>5</rate>
</basic>
```

```
<basic>
  <vocab>palatable</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>roomy</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>delicious</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>exceptional</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>wonderful</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>enjoyable</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>enjoy</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>illy</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>ill</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>bustling</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>crowded</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>clear</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>inspiring</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>uninspiring</vocab>
  <rate>2</rate>
</basic>
```

```
<basic>
  <vocab>proper</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>eager</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>lamentable</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>lamentably</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>personable</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>handy</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>near</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>nearest</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>large</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>spacious</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>big</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>huge</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>spacious</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>enormous</vocab>
  <rate>4</rate>
</basic>
```

```
<basic>
  <vocab>small</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>narrow</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>lacking</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>daft</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>scary</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>vast</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>impressive</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>outstanding</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>empty</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>sad</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>tepid</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>aged</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>top</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>low</vocab>
  <rate>1</rate>
</basic>
```

<basic>
 <vocab>ordinary</vocab>
 <rate>3</rate>
</basic>
<basic>
 <vocab>accommodating</vocab>
 <rate>4</rate>
</basic>
<basic>
 <vocab>admirable</vocab>
 <rate>4</rate>
</basic>
<basic>
 <vocab>adorable</vocab>
 <rate>4</rate>
</basic>
<basic>
 <vocab>advisable</vocab>
 <rate>3</rate>
</basic>
<basic>
 <vocab>aesthetic</vocab>
 <rate>4</rate>
</basic>
<basic>
 <vocab>afraid</vocab>
 <rate>2</rate>
</basic>
<basic>
 <vocab>fearful</vocab>
 <rate>2</rate>
</basic>
<basic>
 <vocab>aggressive</vocab>
 <rate>2</rate>
</basic>
<basic>
 <vocab>aidful</vocab>
 <rate>4</rate>
</basic>
<basic>
 <vocab>aimful</vocab>
 <rate>4</rate>
</basic>
<basic>
 <vocab>airy</vocab>
 <rate>4</rate>
</basic>
<basic>
 <vocab>allergic</vocab>
 <rate>2</rate>
</basic>
<basic>
 <vocab>annoying</vocab>
 <rate>2</rate>
</basic>

```
<basic>
  <vocab>anxious</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>awesome</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>awkward</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>base</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>bland</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>brilliant</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>careful</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>attentive</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>alright</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>ripoff</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>barely</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>highest</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>bottom</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>easy</vocab>
  <rate>1</rate>
</basic>
```

```
<basic>
  <vocab>effective</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>working</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>work</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>dire</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>poor</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>creak</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>grumpy</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>slippery</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>simple</vocab>
  <rate>3</rate>
</basic>
<basic>
  <vocab>tidy</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>horrible</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>trendy</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>sloppy</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>ignorant</vocab>
  <rate>2</rate>
</basic>
```

```
<basic>
  <vocab>sick</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>filthy</vocab>
  <rate>1</rate>
</basic>
<basic>
  <vocab>elegant</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>cosy</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>amazing</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>slow</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>fast</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>bountiful</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>special</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>much</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>extraordinary</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>courteous</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>harsh</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>ritzy</vocab>
  <rate>4</rate>
</basic>
```



```

<basic>
  <vocab>luxurious</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>useful</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>comfy</vocab>
  <rate>4</rate>
</basic>
<basic>
  <vocab>steep</vocab>
  <rate>2</rate>
</basic>
<basic>
  <vocab>terrific</vocab>
  <rate>5</rate>
</basic>
<basic>
  <vocab>faultless</vocab>
  <rate>5</rate>
</basic>
</basics>
<advances>
  <advance>
    <vocab>very</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>
  <advance>
    <vocab>incredibly</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>
  <advance>
    <vocab>somewhat</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>
  <advance>
    <vocab>super</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>
  <advance>
    <vocab>so</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>
  <advance>
    <vocab>abundantly</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>

```

```

<advance>
  <vocab>really</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>highly</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>totally</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>extremely</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>more</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>below</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>-1</ratemoreequal>
</advance>
<advance>
  <vocab>most</vocab>
  <rateless>-2</rateless>
  <ratemoreequal>+2</ratemoreequal>
</advance>
<advance>
  <vocab>definitely</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>quite</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>absolutely</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>
  <vocab>certainly</vocab>
  <rateless>-1</rateless>
  <ratemoreequal>+1</ratemoreequal>
</advance>
<advance>

```

```

    <vocab>perfectly</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>
  <advance>
    <vocab>deep</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>+1</ratemoreequal>
  </advance>
</advances>
<specials>
  <special>
    <vocab>too</vocab>
    <rateless>-1</rateless>
    <ratemoreequal>-1</ratemoreequal>
  </special>
  <special>
    <vocab>nothing</vocab>
    <rateless>+2</rateless>
    <ratemoreequal>-2</ratemoreequal>
  </special>
  <special>
    <vocab>never</vocab>
    <rateless>+2</rateless>
    <ratemoreequal>-2</ratemoreequal>
  </special>
  <special>
    <vocab>not</vocab>
    <rateless>+2</rateless>
    <ratemoreequal>-2</ratemoreequal>
  </special>
  <special>
    <vocab>pretty</vocab>
    <rateless>+1</rateless>
    <ratemoreequal>-1</ratemoreequal>
  </special>
  <special>
    <vocab>relatively</vocab>
    <rateless>+1</rateless>
    <ratemoreequal>-1</ratemoreequal>
  </special>
  <special>
    <vocab>almost</vocab>
    <rateless>+1</rateless>
    <ratemoreequal>-1</ratemoreequal>
  </special>
  <special>
    <vocab>abit</vocab>
    <rateless>+1</rateless>
    <ratemoreequal>-1</ratemoreequal>
  </special>
</specials>
</words>


```

ภาพที่ ก.1 คลังคำศัพท์ของระบบ

2. ตัวอย่างข้อมูลที่ใช้ในการทดสอบ

ในการประเมินความถูกต้องของการสกัดความรู้ได้นำตัวอย่างบทวิจารณ์ที่มีระดับความพึงพอใจโดยรวมแตกต่างกัน 5 ระดับ คือ ระดับที่ 1-5 มาทดสอบ ดังแสดงในรูปแบบตัวอย่างที่ ก.2-ก.6 บทวิจารณ์แต่ละบทจะประกอบไปด้วยคำวิจารณ์ที่กล่าวถึงเรื่องต่าง ๆ ซึ่งมีทั้งที่เขียนถูกและผิดไวยากรณ์รวมทั้งกล่าวถึงเรื่องที่เกี่ยวข้องและไม่เกี่ยวข้องกับการสกัดความรู้

[filthy](#)

Traveler rating:  2 | 3 | 4 | 5 | **1.0**

Jul 17, 2008

We booked this hotel on the basis. it had a family room, 1 double 2 singles. It was on the top floor with a lift for 1 person. The room and bathroom were so bad. we considered booking in to the closest holiday inn and losing our money but put up with it because it was only for 1 night. The breakfast was a joke. They even watered down the milk!!

This TripAdvisor Member

Liked: Nothing

Disliked: Cleanliness

Posted by [hardybandit](#), birmingham

This review is the subjective opinion of a TripAdvisor member and not of TripAdvisor LLC.

ภาพที่ ก.2 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 1

Not great

Traveler rating:  2.0

May 31, 2005

The room was dark, stale-smelling and untidy (curtains hanging off rail) and there were NO tea/coffee making facilities at all in the room despite publicity stating that "every one of the 120 rooms has...."

The room was in the basement and had a glass door leading straight outside (a security and privacy risk in my view).

The bed was extremely uncomfortable, the television reception was poor.

Sound insulation was good, no noise from neighbours at all.

The shower and toilet were efficient, but there were no drinking glasses provided at all (for such essentials as rinsing after teeth-brushing !)

Not a totally bad experience (I've had much worse) but a little effort had been spared and the lack of tea/coffee facilities (in 2005) was unacceptable.

Posted by [Factor50](#), UK

This review is the subjective opinion of a TripAdvisor member and not of TripAdvisor LLC.

ok for people on a budget

Traveler rating:  3.0

Jun 20, 2004

I stayed here for a week whilst working in London. I was pleased to find a hotel for £50/night for a double room including breakfast which was apparently 1/3 of the normal cost of £140/night (through laterooms.com). The room was extremely small. It had tea/coffee facilities. Continental breakfast consisted of croissants, jam & a choice of 2 cereals. Hot food was extra cost.

On the plus point, the location is fantastic. Not bad for people on a budget. I had no problem with noise & the place was clean & rooms daily serviced.

Posted by [A TripAdvisor Member](#)

This review is the subjective opinion of a TripAdvisor member and not of TripAdvisor LLC.

ภาพที่ ก.4 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 3

Cozy place to stay**Traveler rating:**  1 2 3 4 5 4.0

Oct 10, 2008

My husband and I spent three nights at the Rhodes. Chris the owner was a wonder host. Although the room was small it was very very clean and very comfortable. The location was excellent, within walking distance to main roads, Hyde Park and Paddington station. An excellent value!

This TripAdvisor Member**Liked:** Location**Disliked:** SmallPosted by [LeslieL](#), Grimsby Ontario*This review is the subjective opinion of a TripAdvisor member and not of TripAdvisor LLC.*

ภาพที่ ก.5 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 4



[a pearl](#)

Traveler rating:  5.0

Oct 28, 2009

relaxing atmosphere... room not very roomy but very clean and well furnished.

I recommend. Staff very welcoming and helpful.

Posted by [anah_wassermann](#), Tel Aviv

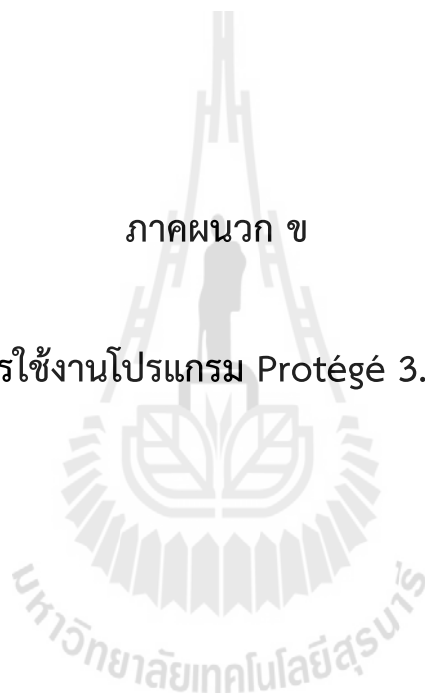
This review is the subjective opinion of a TripAdvisor member and not of TripAdvisor LLC.

ภาพที่ ก.6 ตัวอย่างบทวิจารณ์ในระดับความพึงพอใจที่ 5

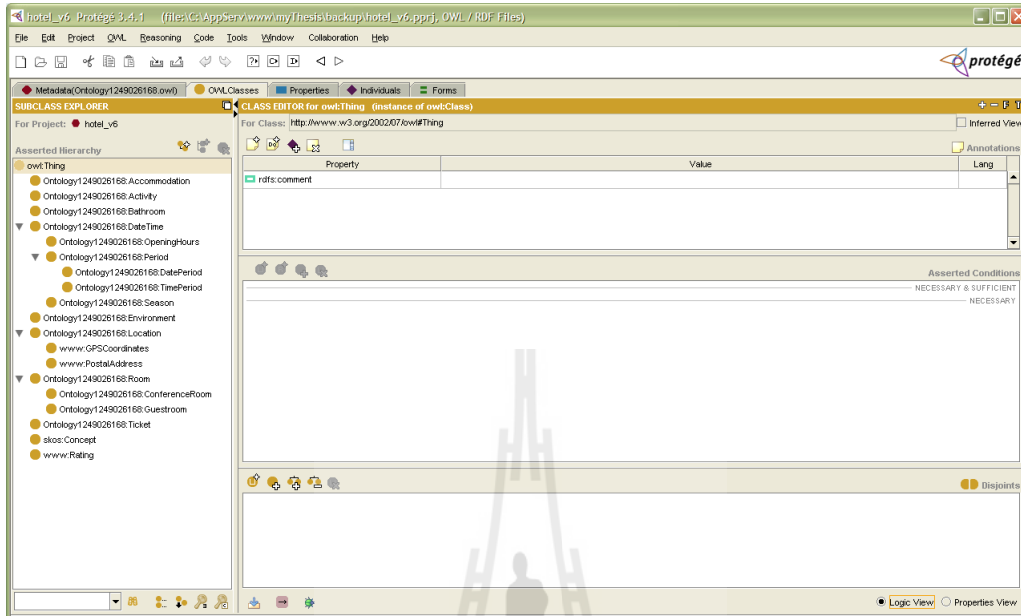


ภาคผนวก ข

การใช้งานโปรแกรม Protégé 3.4.1



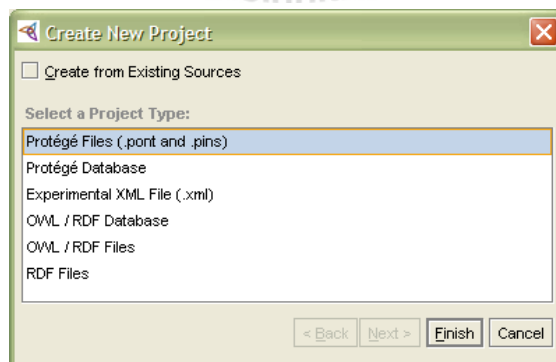
ในส่วนนี้ ครอบคลุมการใช้งานโปรแกรม Protégé เรื่องการสร้างและบันทึกโครงการ การสร้างคลาส การสร้างสล็อต และการสร้างอินสแตนซ์ ซึ่งมีรูปแบบและการทำงาน ดังต่อไปนี้



ภาพที่ ข.1 หน้าจอหลักของโปรแกรม Protégé

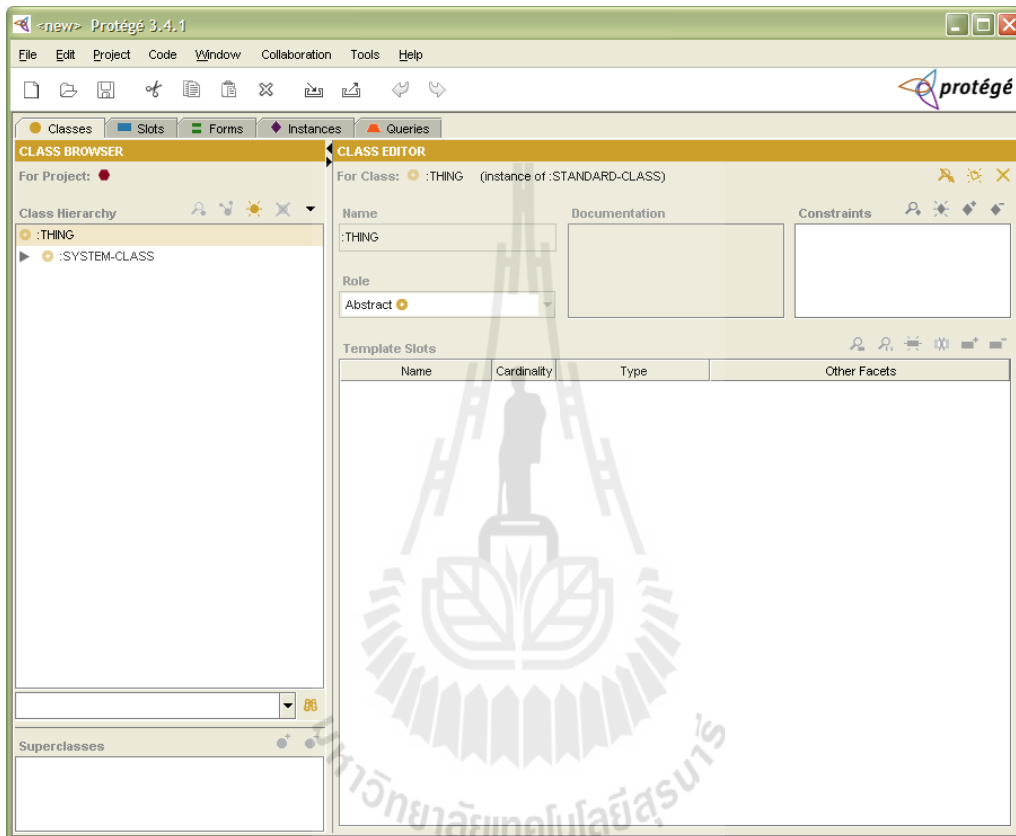
1) การสร้างโครงการ มีขั้นตอนดังนี้

- เมื่อเปิดโปรแกรม Protégé จะปรากฏกล่องข้อความโต้ตอบกับผู้ใช้ ดังภาพที่ ข.2 เพื่อสร้างโครงการใหม่, เปิดโครงการที่มีอยู่แล้ว หรือดูข้อความช่วยเหลือ



ภาพที่ ข.2 กล่องข้อความโต้ตอบกับผู้ใช้เพื่อสร้างหรือเปิดโครงการ

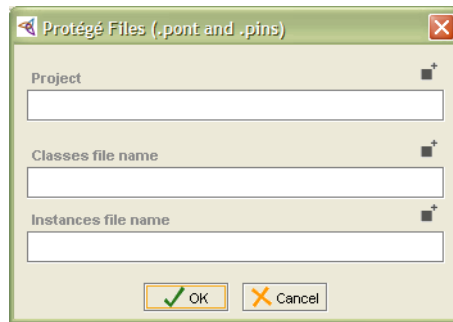
- ในพื้นที่ “Project Format” ด้านซ้ายมือ เลือก Default Files (.pont and pins) เพื่อสร้างไฟล์โครงงานในรูปแบบมาตรฐานของ Protégé
- คลิกที่ปุ่ม New จะปรากฏหน้าต่างของโปรแกรม Protégé และเห็นมาตรฐานต่าง ๆ ในการทำงาน ซึ่งในแท็บคลาสจะบรรจุคลาส THING และ SYSTEM-CLASS ดังแสดงในภาพที่ ข.3



ภาพที่ ข.3 โปรแกรม Protégé แสดงรายละเอียดของแท็บคลาส

2) การบันทึกโครงงาน มีขั้นตอนดังนี้

- เลือกรายการ “Save Project” จากเมนู File จะปรากฏกล่องข้อความ ดังภาพที่ ข.4

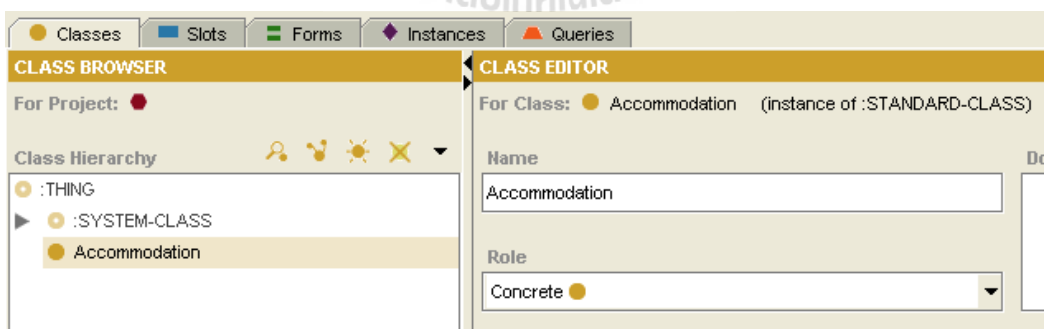


ภาพที่ ข.4 กล่องข้อความโต้ตอบกับผู้ใช้เพื่อบันทึกโครงการ

- เลือกตำแหน่งที่จะบันทึกโครงการโดยคลิกที่ปุ่มด้านขวาบนของกล่องข้อความ Project
- พิมพ์ชื่อโครงการที่ต้องการบันทึก เช่น e-Tourism
- คลิกปุ่ม OK

3) การสร้างคลาส มีขั้นตอนดังนี้

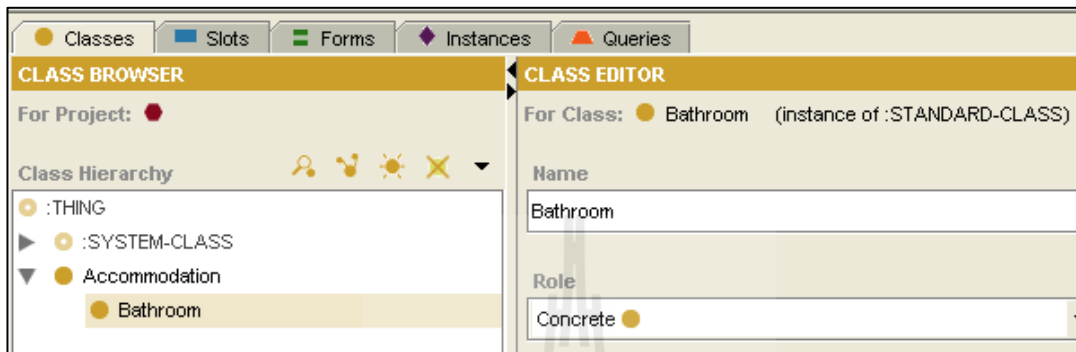
- ที่แท็บคลาส คลิกเมาส์ขวาที่คลาส THING จะปรากฏเมนู Pop up ให้เลือกเมนู “Create Class” หรือ คลิกเลือกที่คลาส THING แล้วกดปุ่ม  (Create Class) ซึ่งอยู่ที่ด้านบนขวา
- กำหนดชื่อคลาส ในช่อง Name เป็นค่า “Accommodation”
- กด Enter หรือ คลิกที่ชื่อคลาสที่ถูกเน้น เพื่อให้ชื่อคลาสที่เปลี่ยนแปลงไปปรากฏ



ภาพที่ ข.5 หน้าจอแสดงการสร้างคลาส

การสร้างคลาสย่อย (Subclass) ของคลาส Accommodation

- คลิกเลือกคลาส Accommodation
- คลิกปุ่ม Create Class
- ตั้งชื่อคลาสย่อย เป็น “Bathroom”




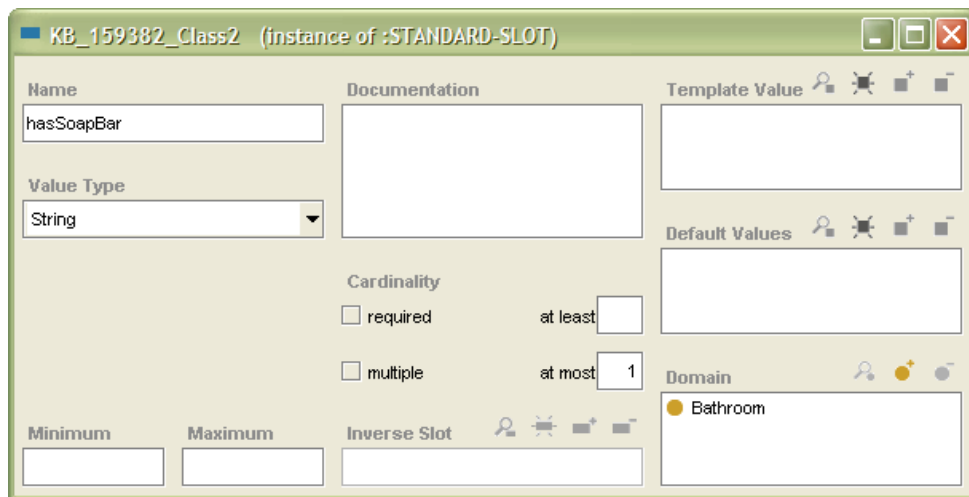
ภาพที่ ข.6 หน้าจอแสดงการสร้างคลาสย่อย

เมื่อได้สร้างคลาสย่อยอันแรกให้กับคลาสใด ๆ จะปรากฏเครื่องหมาย ▶ หรือ ▼ ที่ด้านซ้ายของคลาสซึ่งใช้สำหรับการแสดงและซ่อนคลาสย่อยของคลาสนั้น ๆ ตามลำดับ

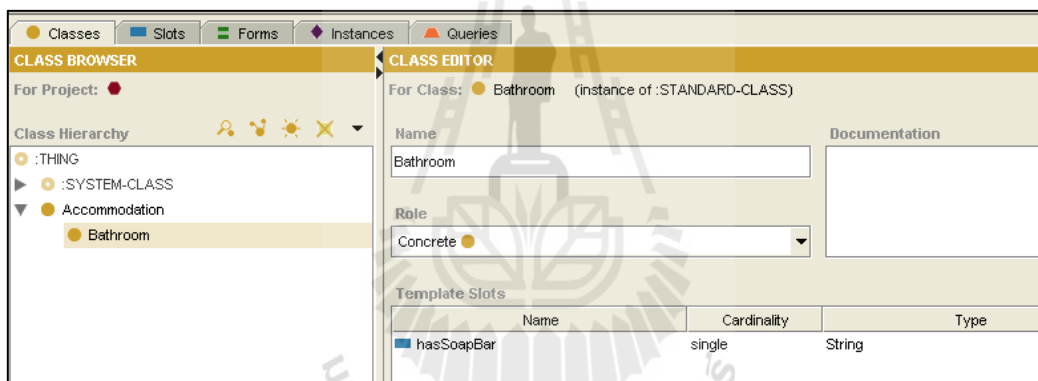
4) การสร้างสล็อต

เป็นการกำหนดคุณสมบัติของคลาสหรือกำหนดความสัมพันธ์ (Relation) ระหว่างคลาส ในที่นี้จะกำหนดสล็อตต่าง ๆ ให้กับคลาส “Bathroom” คือ คุณสมบัติ “hasSoapBar” และ กำหนดความสัมพันธ์ “hasRating” ให้ระหว่างคลาส “Rating” และคลาส “Bathroom” ดังนี้


- เลือกคลาส “Bathroom”
- กดปุ่ม  (Create Slot) ที่ส่วนข้อมูลแม่แบบสล็อต (Template Slot) จะปรากฏสล็อตฟอร์ม (Slot Form) ดังภาพที่ ข.7
- เปลี่ยนชื่อสล็อตในกล่องข้อความ Name เป็น “hasSoapBar” และกำหนดค่า Value Type หรือ ชนิดของข้อมูลของคุณสมบัตินี้ เป็น String
- จะปรากฏสล็อต “hasSoapBar” ในส่วนข้อมูลแม่แบบสล็อต ดังภาพที่ ข.8

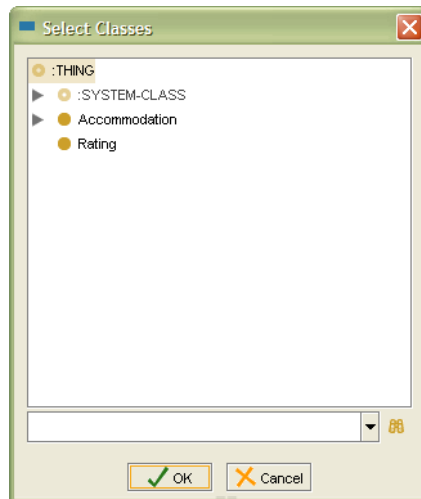


ภาพที่ ข.7 หน้าต่างสล็อตฟอร์ม

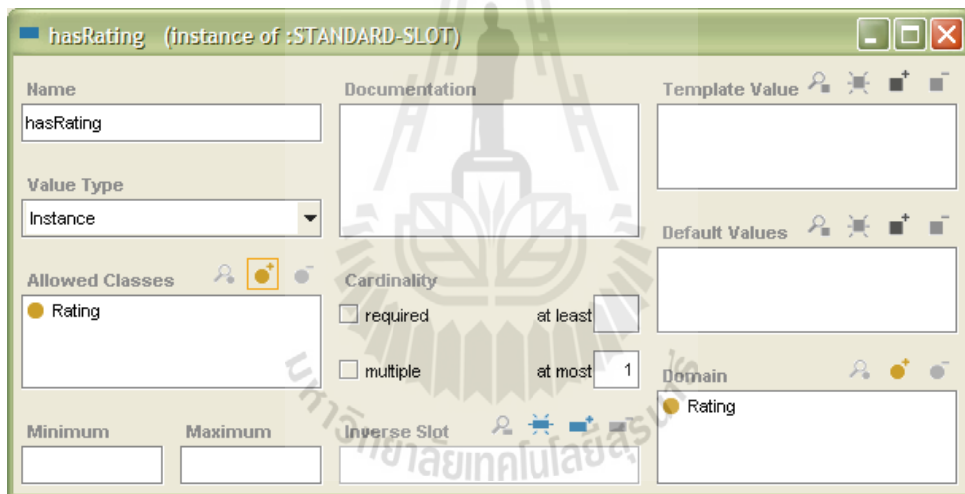


ภาพที่ ข.8 หน้าจอแสดงรายละเอียดสล็อตในส่วนของแม่แบบสล็อต

- สร้างสล็อต “hasRating” โดยกำหนดค่า Value Type ให้กำหนดเป็น Instance และที่ส่วนข้อมูล Allowed Classes คลิกปุ่ม  (Add Class) จะปรากฏหน้าต่างดังภาพที่ ข.9 ให้ผู้ใช้เลือกคลาสที่จะมาทำความสัมพันธ์ หรือเป็นค่า Range ของความสัมพันธ์นี้ ให้เลือกคลาส “Rating” จากนั้นจะปรากฏคลาส Rating ในส่วนข้อมูล Allowed Classes ดังภาพที่ ข.10




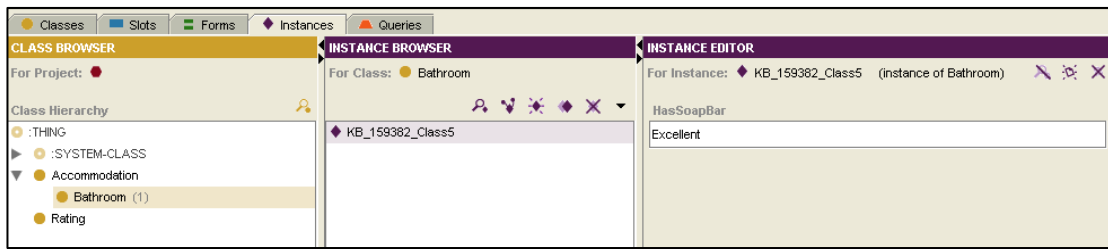
ภาพที่ ข.9 หน้าต่างเลือกคลาสที่จะนำมาสร้างความสัมพันธ์



ภาพที่ ข.10 ผลการสร้างความสัมพันธ์ระหว่างคลาส

5) การสร้างอินสแตนซ์

- คลิกที่แท็บอินสแตนซ์และคลิกเลือกคลาส “Bathroom”
- ในส่วนข้อมูล INSTANCE BROWSER คลิกปุ่ม  (Create Instance) แล้วใส่ข้อมูลในส่วนของ INSTANCE EDITOR ดังภาพที่ ข.11



ภาพที่ ข.11 ผลการสร้างอินสแตนซ์



ประวัติผู้วิจัย

นางจิตติมนต์ อึ้งสกุล ตำแหน่งอาจารย์ประจำสาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี เกิดวันที่ ๓ พฤศจิกายน พ.ศ. ๒๕๒๑ ที่จังหวัดกรุงเทพมหานคร จบการศึกษาระดับปริญญาตรี ปริญญาโท และปริญญาเอก สาขาวิศวกรรมคอมพิวเตอร์ จากมหาวิทยาลัยเกษตรศาสตร์ ในปีพ.ศ. ๒๕๔๒ พ.ศ. ๒๕๔๔ และ พ.ศ. ๒๕๔๘ ตามลำดับ ปัจจุบันได้ตีพิมพ์เผยแพร่ผลงานวิจัยในวารสารและงานประชุมทางวิชาการทั้งระดับชาติและนานาชาติรวม ๔๑ เรื่อง สามารถติดต่อได้ที่ ๑๑๑ ถ. มหาวิทยาลัย ต. สุรนารี อ. เมืองนครราชสีมา จ. นครราชสีมา ๓๐๐๐๐

นายธรา อึ้งสกุล ตำแหน่งอาจารย์ประจำสาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี เกิดวันที่ ๑๖ พฤศจิกายน พ.ศ. ๒๕๒๑ ที่จังหวัดกรุงเทพมหานคร จบการศึกษาระดับปริญญาตรี และปริญญาโท สาขาวิศวกรรมคอมพิวเตอร์ จากมหาวิทยาลัยเกษตรศาสตร์ ในปี พ.ศ. ๒๕๔๒ และ พ.ศ. ๒๕๔๕ ตามลำดับ และจบการศึกษาระดับปริญญาเอก สาขาวิทยาการคอมพิวเตอร์ จากมหาวิทยาลัยแห่งรัฐเทนเนสซี ประเทศสหรัฐอเมริกา ในปีพ.ศ. ๒๕๕๐ ปัจจุบันได้ตีพิมพ์เผยแพร่ผลงานวิจัยในวารสารและงานประชุมทางวิชาการทั้งระดับชาติและนานาชาติรวม ๖๑ เรื่อง สามารถติดต่อได้ที่ ๑๑๑ ถ. มหาวิทยาลัย ต. สุรนารี อ. เมืองนครราชสีมา จ. นครราชสีมา ๓๐๐๐๐