บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาวิจัยครั้งนี้ผู้วิจัยได้นำเสนอแนวความคิด ทฤษฎีและได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ซึ่งมีรายละเอียดดังนี้

1. แนวคิดเกี่ยวกับการท่องเที่ยว
2. หลักการและเครื่องมือออนโทโลจี (Ontology and Tools)
3. หลักการออนโทโลจีเชิงเวลา (Temporal Ontology)
4. ทฤษฎีเนมแมทชิ่ง (Name Matching)
5. อัลกอริทึมในการจัดหมวดหมู่และแนะนำข้อมูลท่องเที่ยว (Classification Algorithm and Tourism Recommendation)
6. งานวิจัยที่เกี่ยวข้อง

แนวคิดเกี่ยวกับการท่องเที่ยว

สุรางค์รัตน์ เชาว์โคกสูง (2554) ได้อธิบายแนวคิดเกี่ยวกับการท่องเที่ยวว่าเป็นเรื่องของการเดินทางหรือการวางแผนเดินทางจากสถานที่หนึ่งไปอีกสถานที่หนึ่ง โดยใช้ยานพาหนะนำไป การเดินทางท่องเที่ยวนั้นหมายถึง การเดินทางที่มีเงื่อนไข 3 ประการ คือ

1. การเดินทางจากที่อยู่อาศัยปกติไปยังที่อื่นเป็นการชั่วคราว
2. การเดินทางด้วยความสมัครใจ
3. การเดินทางด้วยวัตถุประสงค์ใดๆ ก็ตามที่ไม่ใช่เป็นการประกอบอาชีพหรือหารายได้

ดังนั้นการเดินทางจึงมิได้จำกัดเฉพาะเพียงการเดินทางเพื่อพักผ่อนหย่อนใจ หรือ เพื่อความสนุกสนาน บันเทิงเริงรมย์ ดังที่คนส่วนมากเข้าใจกัน การเดินทางเพื่อการประชุมสัมมนา เพื่อศึกษาหาความรู้ เพื่อการกีฬา เพื่อการติดต่อธุรกิจ ตลอดจนการเยี่ยมเยียนญาติพี่น้อง นับว่าเป็นการท่องเที่ยวทั้งสิ้น ในการเดินทางเพื่อการท่องเที่ยวนั้น มีเหตุในการจูงใจหรือมุ่งมั่นในการเดินทาง ดังนี้

1. เพื่อความสนุกสนานเพลิดเพลิน การทำงานหนักซ้ำซากจะทำให้เกิดความเบื่อหน่าย ดังนั้นคนเราจึงต้องการเปลี่ยนแปลงบรรยากาศเพื่อหาความสนุกสนานเพลิดเพลิน

2. เพื่อการพักผ่อน หลังจากทำงานหนัก คนเราต้องการการผ่อนคลายความเครียดจากการทำงาน ดังนั้นโปรแกรมท่องเที่ยวเพื่อการพักผ่อนจึงมักเป็นที่ที่มีอากาศดี ภูมิประเทศสวยงาม สงบเงียบและมีสิ่งเจริญตาเจริญใจหัวข้อสำคัญ

3. การท่องเที่ยวเชิงวัฒนธรรม สถานที่ของนักท่องเที่ยวกลุ่มนี้จะได้แก่ โบราณสถาน โบราณวัตถุหรือสถานที่ที่มีประเพณี วัฒนธรรมของตนเอง

4. เพื่อการกีฬา แบ่งออกเป็นเพื่อการชมกีฬา และเล่นกีฬา เช่น การแข่งรถ ปีนเขา เล่นเรือใบ เป็นต้น

5. เพื่อธุรกิจ หมายถึง การที่ไปทำธุรกิจและท่องเที่ยวไปด้วย

6. เพื่อประชุมสัมมนา ปัจจุบันการจัดประชุมสัมมนา มักนิยมจัดตามที่ต่างๆ รวมทั้งต่างประเทศ จึงเป็นเหตุให้ผู้เข้าร่วมประชุมมีโอกาสท่องเที่ยวในประเทศหรือท้องถิ่นนั้นๆ ได้เช่นกัน

7. อาชีพ ภารกิจการงานสำหรับบางคนอาจเป็นอุปสรรคในการเดินทาง

8. ระยะทาง การเดินทางไปยังสถานที่ท่องเที่ยวที่ไกล อาจจะไม่เอื้ออำนวยและสมดุลกับเวลาว่าง การเงิน หรือสุขภาพด้วย

เดือนรัตน์ รัตนศิริ (2553 อ้างอิงใน สุรางค์รัตน์ เชาว์โคกสูง, 2554) ได้ให้แนวคิดเกี่ยวกับสิ่งกระตุ้น จูงใจให้นักท่องเที่ยวเดินทางมาท่องเที่ยวหรือเรียกว่าสินมรดกทางการท่องเที่ยวว่ามี 2 ประการ คือ

1. สินมรดกตามธรรมชาติ ได้แก่ ลักษณะภูมิประเทศ สัตว์ ป่า ภูเขา ทะเล น้ำตก เป็นต้น

2. สินมรดกที่มนุษย์สร้างขึ้น เพื่อจูงใจให้เกิดการท่องเที่ยวและเพื่อให้การท่องเที่ยวเป็นไปอย่างสะดวกสบาย ได้แก่

2.1 สินมรดกทางการท่องเที่ยวในเชิงประวัติศาสตร์ วัฒนธรรมและศาสนา

2.2 โครงสร้างพื้นฐานทางการท่องเที่ยว ได้แก่ ระบบสาธารณูปโภคต่างๆ เครื่องอำนวยความสะดวกอื่นๆ เพื่อให้การเดินทางไปยังแหล่งท่องเที่ยวสะดวกและปลอดภัย

2.3 โครงสร้างเหนือฐานรอง ได้แก่ สิ่งที่ช่วยดึงดูดนักท่องเที่ยวให้มาท่องเที่ยวมากขึ้น เช่น โรงแรม ภัตตาคาร ร้านขายของที่ระลึก

2.4 วิถีการดำรงชีวิตของคน ซึ่งรวมถึงท่าที ความรู้สึกของประชาชนในแหล่งท่องเที่ยว เช่น ความยิ้มแย้มแจ่มใส ความมีน้ำใจ ความพร้อมที่จะช่วยเหลือนักท่องเที่ยว

นอกจากนี้ การท่องเที่ยวแห่งประเทศไทย (2540) ได้กล่าวถึง การรับรู้ถึงการจัดกิจกรรมและบริการสิ่งอำนวยความสะดวกประเภทที่เกี่ยวข้องกับนักท่องเที่ยวไว้ 10 ประเภท คือ

1. การคมนาคมขนส่ง
2. ที่พักแรม
3. ร้านอาหารและภัตตาคาร
4. บริการนำเที่ยวและมัคคุเทศก์
5. สิ่งที่ดึงดูดใจทางการท่องเที่ยว
6. ร้านขายของที่ระลึกและขายสินค้าพื้นเมือง
7. ความปลอดภัย
8. การอำนวยความสะดวกในการเข้า–ออกเมือง
9. การจัดบริการอื่นๆ
10. การเผยแพร่โฆษณา

บุญเลิศ จิตตั้งวัฒนา (2548) อ้างอิงจากการท่องเที่ยวแห่งประเทศไทยในการแบ่งประเภทของทรัพยากรท่องเที่ยวออกเป็น 3 ประเภท คือ

1. ทรัพยากรท่องเที่ยวประเภทธรรมชาติ เป็นสถานที่ท่องเที่ยวที่เกิดขึ้นเองตามธรรมชาติ โดยเกิดจากสภาพแวดล้อมทางกายภาพตามธรรมชาติที่มีความสวยงามน่าสนใจต่อการเดินทางไปเที่ยวชม ซึ่งรวมถึงบริเวณที่มนุษย์เข้าไปปรุงแต่งเพิ่มเติมจากสภาพธรรมชาติในบางส่วน เช่น ป่า เขา เกาะ แก่ง ชายหาด ถ้ำ ทะเล แม่น้ำ ทะเลสาบ อ่างเก็บน้ำ กว๊าน น้ำตก น้ำพุร้อน ทุ่งหญ้า ทุ่งดอกไม้ ลานหิน เนินดิน เสาหิน ธารน้ำแข็ง สัตว์ป่า เป็นต้น

2. ทรัพยากรท่องเที่ยวประเภทประวัติศาสตร์ โบราณสถานและโบราณวัตถุ เป็นทรัพยากรท่องเที่ยวที่มนุษย์สร้างขึ้นตามประโยชน์ของมนุษย์เองทั้งที่เป็นมรดกในอดีตและได้สร้างเสริมขึ้นในปัจจุบัน ซึ่งมีความสำคัญทางประวัติศาสตร์ โบราณคดีและศาสนา เช่น ศาสนสถาน อุทยานประวัติศาสตร์ กำแพงเมือง คูเมือง เมืองโบราณ อนุสาวรีย์ อนุสรณ์สถาน ชุมชนโบราณ สถาปัตยกรรมโบราณ ภาพเขียนโบราณ โบราณวัตถุ สมบัติโบราณ พระพุทธรูปโบราณ รูปปั้นโบราณ เป็นต้น

3. ทรัพยากรท่องเที่ยวประเภทศิลปวัฒนธรรม ประเพณีและกิจกรรม เป็นทรัพยากรท่องเที่ยวที่มนุษย์สร้างขึ้นในรูปแบบของการดำเนินชีวิตของผู้คนในสังคมแต่ละกลุ่มชนที่มีความแตกต่างกันไปตามสภาพแวดล้อมทางธรรมชาติ โดยยึดถือปฏิบัติสืบทอดต่อกันมาตลอดจนกิจกรรมต่างๆ ของสังคมที่มีผลต่อการดึงดูดให้นักท่องเที่ยวเดินทางไปเยือน เช่น ชุมชน หมู่บ้าน เรือนแพ ตลาด ตลาดน้ำ ศูนย์วัฒนธรรม การละเล่นพื้นบ้าน ดนตรีพื้นเมือง สินค้าพื้นเมือง ไร่สวน เหมือง วิถีชีวิต อัธยาศัยไมตรีของประชาชน ประเพณีสงกรานต์ ประเพณีแห่เทียนพรรษา เทศกาล

กินเจ ประเพณีลอยกระทง งานบุญบั้งไฟ งานพิธีโยนบัว ประเพณีอินทขิล กิจกรรมแข่งขันกีฬา กิจกรรมบันเทิง สวนสนุก สวนสาธารณะเฉพาะทาง สนามกอล์ฟ สนามแข่งรถ เขื่อนพลังไฟฟ้า กิจกรรมการเดินป่า ดูนก ส่องสัตว์ การไต่เขา การตกปลา การแสวงบุญ การพายเรือ การดำน้ำ การล่องแพ การตั้งแค้มป์ เป็นต้น

บุญเลิศ จิตตั้งวัฒนา (2555) ได้ศึกษาพฤติกรรมและความต้องการของกลุ่มตลาดนักท่องเที่ยวตามวัย ซึ่งสอดคล้องกับช่วงอายุ (สารานุกรมไทยสำหรับเยาวชนฯ, 2530) ดังต่อไปนี้

1. วัยเด็ก ช่วงอายุระหว่างแรกเกิด - 12 ปี จะเป็นกลุ่มที่มีอิทธิพลต่อการตัดสินใจท่องเที่ยวของครอบครัวเป็นอย่างมาก สถานที่ท่องเที่ยวหรือกิจกรรมที่กลุ่มนี้สนใจจะต้องมีความสนุกสนานเป็นสำคัญ เช่น การไปเที่ยวสวนสนุก การเข้าค่ายเยาวชน การไปชมการแสดงโชว์ต่างๆ ภาพยนตร์การ์ตูน การไปเที่ยวสวนสัตว์ เป็นต้น

2. วัยรุ่น ช่วงอายุระหว่าง 13 - 18 ปี จะชื่นชอบการท่องเที่ยวแบบพักผ่อนหย่อนใจตามสมัย สถานที่ท่องเที่ยวหรือกิจกรรมที่กลุ่มนี้สนใจ ได้แก่ การเดินทางและซื้อของในห้างสรรพสินค้า ท่องเที่ยวกลางคืน สังสรรค์กับเพื่อนฝูง เป็นต้น

3. วัยหนุ่มสาว ช่วงอายุระหว่าง 19 - 25 ปี หากเป็นหนุ่มสาวที่สมรสแต่ยังไม่มีบุตรจะเรียกว่าอยู่ในระยะ "น้ำผึ้งพระจันทร์" สถานที่ท่องเที่ยวหรือกิจกรรมที่กลุ่มนี้สนใจมักจะเป็นสถานที่ที่มีสิ่งอำนวยความสะดวกครบครัน บรรยากาศดี โรแมนติก เช่น ทะเล ต่างประเทศ ซื้อของตกแต่งบ้าน เป็นต้น

4. วัยผู้ใหญ่ ช่วงอายุระหว่าง 26 - 35 ปี กลุ่มนี้ส่วนมากจะมีครอบครัวและมีบุตร ดังนั้นการท่องเที่ยวส่วนใหญ่จะเป็นการตามใจลูกๆ โดยจะคำนึงถึงความสะดวกสบายและความปลอดภัยเป็นอันดับแรก เช่น ทะเล สวนสัตว์ สวนสนุก เป็นต้น แต่หากไม่มีบุตรก็มักจะท่องเที่ยวตามใจตนเองที่เน้นการผจญภัย ความตื่นเต้น ความบันเทิง ความสะดวกสบาย เป็นต้น

5. วัยกลางคน ช่วงอายุระหว่าง 36 - 55 ปี กลุ่มนี้จะเหมือนกับวัยผู้ใหญ่แต่จะชอบความสะดวกสบายมากกว่าและความปลอดภัยเป็นอันดับแรก เช่น ทะเล สวนสัตว์ สวนสนุก เป็นต้น แต่หากไม่มีบุตรก็มักจะท่องเที่ยวตามใจตนเองที่เน้นการผจญภัย ความตื่นเต้น ความบันเทิง ความสะดวกสบาย เป็นต้น

6. วัยสูงอายุ อายุตั้งแต่ 55 ปีขึ้นไป กลุ่มนี้จะเป็นกลุ่มที่มีทั้งเวลาและเงิน แต่สุขภาพก็จะเริ่มทรุดโทรมลงไปตามลำดับ พฤติกรรมการท่องเที่ยวของกลุ่มนี้จะต้องมีความสะดวกสบาย มีการบริการดูแลเป็นอย่างดี มีระยะเวลาในการเดินทางไปท่องเที่ยวไม่นานนัก สถานที่ท่องเที่ยวหรือ

กิจกรรมที่กลุ่มนี้สนใจ เช่น ไปเที่ยวชมโบราณสถาน ไปเที่ยวชมวัด ไปทำบุญ เที่ยวชมและเลือกซื้อสินค้าหัตถกรรม เป็นต้น

อรจนา จันทรประยูร (2555) ได้ศึกษาพฤติกรรมและความต้องการของกลุ่มตลาดนักท่องเที่ยวตามเพศ ดังต่อไปนี้

1. เพศหญิง ส่วนใหญ่ชอบศึกษาชีวิตความเป็นอยู่ของคน ทำความรู้จักคน เยี่ยมชมสถานที่ต่างๆ ลองอาหารท้องถิ่น ศึกษาวัฒนธรรม ซื้อของที่ระลึกหรือของเก่า เล่นกีฬาทางบกทางน้ำ อาบแดด ฟังเพลง ชอบใช้ชีวิตในยามราตรีบ้าง

2. เพศชาย ส่วนใหญ่ชอบเที่ยวกลางคืน อาบอบนวด (เพื่อเพศรส) การผจญภัย เดินป่า ดื่ม ซื้อของบ้าง เยี่ยมชมสถานที่ต่างๆ ลองอาหารท้องถิ่นบ้าง

หลักการและเครื่องมือออนโทโลจี (Ontology and Tools)

โดยงานวิจัยนี้ได้นำหลักการออนโทโลจีมาใช้ในการออกแบบและจัดเก็บข้อมูลท่องเที่ยวในประเทศไทยเพื่อให้ข้อมูลท่องเที่ยว โดยใช้โปรแกรม Protégé 3.5 สำหรับออกแบบและจัดเก็บข้อมูลออนโทโลจีท่องเที่ยว ซึ่งใช้ภาษา SPARQL ในการสืบค้นข้อมูลท่องเที่ยวและใช้ Jena API ทำหน้าที่เป็น OWL Parser เพื่ออ่านโครงสร้างของข้อมูลออนโทโลจีท่องเที่ยวที่ถูกเก็บในรูปแบบ OWL ดังรายละเอียดต่อไปนี้

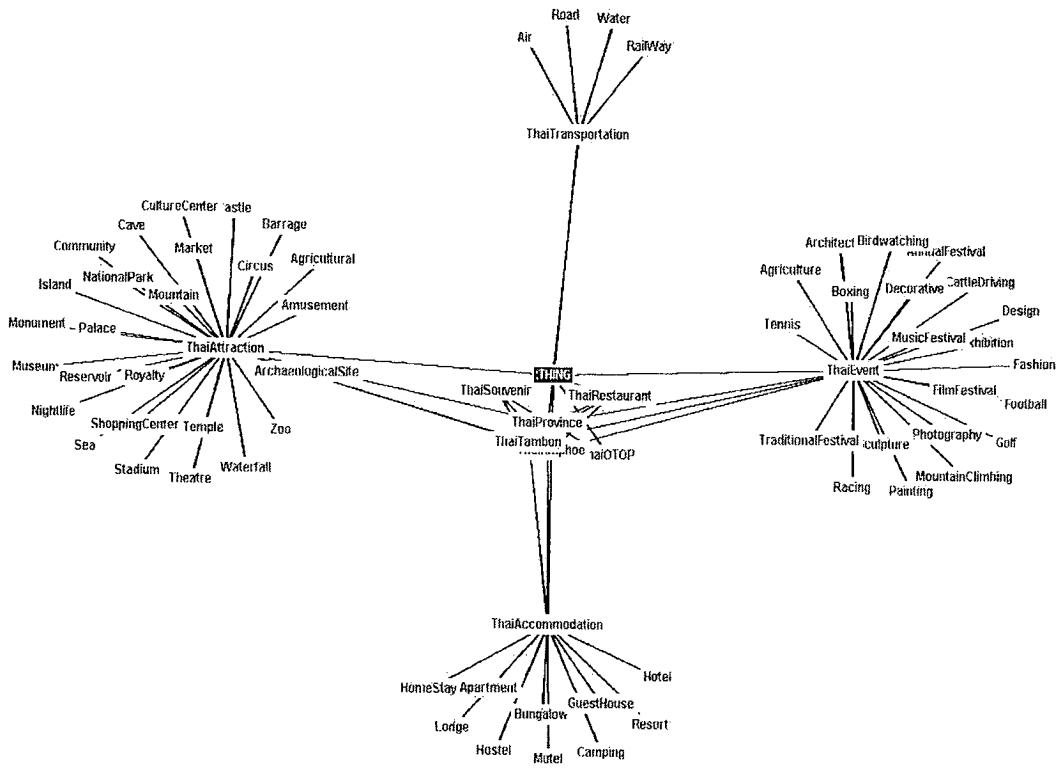
ออนโทโลจี

จักรกฤษณ์ เสน่ห นมะหุต (2557) ได้กล่าวว่าออนโทโลจีได้กลายมาเป็นที่รู้จักและใช้กันอย่างแพร่หลายในปัจจุบัน เช่น สาขาปัญญาประดิษฐ์ (Artificial Intelligence: AI) สาขาวิชาคอมพิวเตอร์ (Computer) และสาขาวิทยาศาสตร์สารสนเทศ (Information Science) โดยเฉพาะอย่างยิ่งในขอบเขตสาขาทางด้านระบบสารสนเทศ (Information System) การรวบรวมสารสนเทศทางปัญญา (Intelligence Information Integration) การสืบค้นข้อมูล (Information Retrieval and Extraction) การแทนความรู้ (Knowledge Representation) การออกแบบฐานข้อมูล (Database Design) การจัดการพื้นฐานองค์ความรู้ (Knowledge Base Management) และระบบการจัดการฐานข้อมูล (Database Management System) ในสาขาทั้งหมดที่กล่าวมานั้น ความหมายของออนโทโลจีก็ยังแตกต่างกันอีกด้วย และไม่มีนิยามเฉพาะตัว เช่น Gruber (1993) ได้นิยามความหมายของออนโทโลจีซึ่งเป็นที่รู้จักกันอย่างกว้างขวางมากที่สุดว่าออนโทโลจีเป็นรายละเอียดที่ชัดเจนแน่นอนของแนวความคิด ดังนั้นเมื่อต้องการแสดงหรือระบุแนวความคิดของวัตถุหรือสิ่งของนั้นๆ ควรจะแสดงรายละเอียดที่ชัดเจนและแน่นอน การนิยามนี้อาจจะเกี่ยวข้องกับชื่อของสิ่งที่มีอยู่ในโปรแกรมซึ่งอาจประกอบไปด้วย Classes, Relations, Functions หรือ Objects

ซึ่ง Smith and Welty (2001) ให้ออนโทโลจีเป็นสาขาหนึ่งของวิชาปรัชญาที่จัดการเกี่ยวกับชนิดและโครงสร้างของวัตถุ คุณสมบัติ กระบวนการ ความเกี่ยวข้องในทุกๆ สาขา และขอบเขตของความเป็นจริง แต่ Noy and McGuinness (2001) ให้ออนโทโลจีเป็นรายละเอียดที่ประกาศที่เป็นทางการในรูปแบบของคอนเซ็ปต์ในโดเมนที่สามารถอธิบายได้ด้วยคลาส สล็อต เงื่อนไข และฟาเซต อีกทั้ง Zhang, et al. (2010) ถือว่าออนโทโลจีเป็นการรวมศาสตร์ทางด้านปัญญาประดิษฐ์และภาษาเครื่องเข้าไว้ด้วยกันเพื่อที่จะแบ่งปันความรู้และนำความรู้ต่างๆ กลับมาใช้ได้อีก นอกจากนี้ยังมีการประมวลผลภาษาธรรมชาติ การแทนความรู้ ที่สามารถนำมาใช้ระหว่างพฤติกรรมของมนุษย์และระบบคอมพิวเตอร์ได้ โดยจะก่อให้เกิดประโยชน์ด้านการสืบค้นข้อมูลและการจัดการความรู้

โดยงานวิจัยนี้ได้ใช้หลักการออนโทโลจีออกแบบและจัดเก็บข้อมูลท่องเที่ยวในประเทศไทย (นฤพนธ์ พนาวงศ์ และจักรกฤษณ์ เสน่ห์, 2553ก, 2553ข) ดังภาพ 1 เนื่องจากฐานข้อมูลแสดงข้อมูลต่างๆ ซึ่งเก็บอยู่ในรูปของตารางข้อมูลและมีความสัมพันธ์กันแต่ไม่ได้มีความสัมพันธ์เชิงโครงสร้างต่อกัน การใช้งานเน้นในเรื่องการสืบค้นข้อมูล ทรานเซคชั่น และรายงาน การค้นหาข้อมูลก็จะใช้วิธีการจับคู่คำกันตรงๆ หากมีเงื่อนไขจำเป็นจะต้องป้อนเงื่อนไขลงในฐานข้อมูลให้ครบหรือเขียนโปรแกรมให้รองรับทุกเงื่อนไขที่มี หากมีเงื่อนไขเพิ่มเติมจะต้องเขียนโปรแกรมเพิ่มเติมเพื่อให้รองรับเงื่อนไขที่เพิ่มขึ้น ทำให้เสียเวลา ยุ่งยาก ไม่สะดวกในการใช้งานแบบทันทีทันใด แต่ออนโทโลจีแสดงความรู้จะอยู่ในรูปแบบเชิงโครงสร้างประกอบด้วยคอนเซ็ปต์ที่เชื่อมโยงกันและหากมีข้อมูลๆ นั้นจะเป็นค่าคงที่เรียกว่า Instance โดยมีลักษณะความสัมพันธ์ต่อกันแบบคู่ต่อๆ กันไปเรียกว่า Triple เน้นการบ่งชี้ความหมายที่ซ่อนอยู่จากความสัมพันธ์ของความรู้นั้นๆ การใช้งานออนโทโลจีนั้นจะไม่เน้นไปที่การสืบค้นข้อมูลแต่เป็นการนำไปสู่การตัดสินใจบนฐานความรู้ด้วยวิธีการอนุมาน ซึ่งหากมีเงื่อนไขใหม่เกิดขึ้นไม่จำเป็นจะต้องแก้ไขหรือปรับเปลี่ยนโปรแกรมใดๆ เพียงแต่ป้อนความรู้หรือเพิ่มกฎลงไปในออนโทโลจีเท่านั้น สรุปได้ว่าหากพัฒนาโปรแกรมสืบค้นข้อมูลแบบตรงๆ ก็ควรใช้ฐานข้อมูลแต่หากต้องการสืบค้นเชิงความหมายต้องใช้ออนโทโลจี

จากภาพ 1 แสดงออนโทโลจีท่องเที่ยว ซึ่งแบ่งเป็น 10 คลาสคือ คลาสจังหวัด (Thai Province) คลาสอำเภอ (Thai Amphoe) คลาสตำบล (Thai Tambon) คลาสเทศกาล (Thai Event) คลาสสถานที่ท่องเที่ยว (Thai Attraction) คลาสที่พัก (Thai Accommodation) คลาสการเดินทาง (Thai Transportation) คลาสร้านอาหาร (Thai Restaurant) คลาสร้านขายของฝาก (Thai Souvenir) และคลาสร้านสินค้าหนึ่งผลิตภัณฑ์หนึ่งตำบล (Thai OTOP)



ภาพ 1 แสดงออนโทโลจีท่องเที่ยว

เครื่องมือสำหรับออกแบบและจัดเก็บข้อมูลออนโทโลจี

Norta, et al. (2006) กล่าวว่าเครื่องมือที่ใช้ในการออกแบบและจัดเก็บข้อมูลออนโทโลจีมีจำนวนมาก ซึ่งแต่ละโปรแกรมนั้นจะมีข้อดีและข้อเสียแตกต่างกันไปขึ้นอยู่กับการใช้งานและความสามารถของแต่ละโปรแกรมที่ไม่เหมือนกัน สำหรับงานวิจัยนี้ได้เลือกใช้โปรแกรม Protégé 3.5 ในการออกแบบและจัดเก็บข้อมูลออนโทโลจีท่องเที่ยวในประเทศไทย โดยจัดเก็บในรูปแบบไฟล์ OWL ทั้งนี้ เนื่องจาก Protégé เป็นโปรแกรมที่ใช้งานสะดวกและง่าย อีกทั้งยังเป็นซอฟต์แวร์รหัสเปิด (Open-Source) ที่สามารถใช้งานได้แบบไม่มีค่าใช้จ่าย พัฒนาโดยมหาวิทยาลัยสแตนฟอร์ด (Stanford University) มีส่วนการติดต่อกับผู้ใช้แบบกราฟิกทำให้ใช้งานได้ง่าย มีรูปแบบการทำงานที่รองรับกับผู้ใช้ทั้งแบบบนเครื่องเดียวหรือบนระบบเครือข่าย นอกจากนี้ยังช่วยให้วิศวกรความรู้สามารถถ่ายทอดและจัดเก็บองค์ความรู้ในรูปแบบของออนโทโลจีได้สะดวกและง่ายด้วยการสร้างโดเมนและฐานความรู้ สร้างความสัมพันธ์แบบลำดับชั้น (is-a) รองรับการจัดเก็บข้อมูลได้หลากหลายภาษา รวมถึงความสามารถในการรองรับกลไกการอนุมานที่มีอยู่ในตัวโปรแกรมเองและรองรับการติดตั้งเพิ่มเติมได้ เช่น Jess Inference Engine

ในปัจจุบันมีการนำโปรแกรม Protégé มาใช้ในการออกแบบและจัดเก็บข้อมูลต่างๆ ไม่ว่าจะเป็นข้อมูลทางด้านธุรกิจ เช่น งานวิจัยของ Snae, et al. (2007) และ Snae and Brueckner (2007) ใช้ Protégé ในการออกแบบออนโทโลจีสำหรับจัดเก็บข้อมูลองค์กรธุรกิจของจังหวัดพิษณุโลก ประกอบด้วย ชื่อ ที่อยู่ เวลาทำการ ประเภทธุรกิจ ขนาดของธุรกิจ เพื่อช่วยในการสืบค้นข้อมูลธุรกิจภายในจังหวัดพิษณุโลก สำหรับงานวิจัยที่ใช้ Protégé ในการออกแบบและจัดเก็บข้อมูลท่องเที่ยวนั้นมีจำนวนมาก เช่น งานวิจัยของ Siricharoen (2008) ใช้ Protégé ในการออกแบบออนโทโลจีท่องเที่ยวจากแนวคิดที่หากต้องการไปท่องเที่ยวยังสถานที่หนึ่งนั้นว่าไปแล้วจะมีกิจกรรมอะไรที่น่าสนใจบ้าง สถานที่นั้นเหมาะจะไปช่วงเวลาใด พักที่ไหน การเดินทางไปอย่างไร แบ่งการออกแบบเป็น 4 คลาส คือ Where (เที่ยวที่ไหน พักที่ไหน) When (ไปเมื่อไหร่) What (ทำกิจกรรมอะไร) และ Transportation (เดินทางไปได้อย่างไร) ซึ่ง Ananthapadmanaban and Srivatsa (2011) ใช้ Protégé ในการออกแบบการท่องเที่ยวรัฐทมิฬนาฑู ประเทศอินเดีย โดยแบ่งออนโทโลจีเป็น 2 ส่วน คือ ส่วนของผู้ใช้จัดเก็บข้อมูลนักท่องเที่ยว ประกอบด้วย ลักษณะการเที่ยว (คนเดียว คู่รัก ครอบครัว) จำนวนสมาชิก (เด็กกี่คน ผู้ใหญ่กี่คน) ประเภทสถานที่ท่องเที่ยวที่สนใจ และส่วนข้อมูลท่องเที่ยวรัฐทมิฬนาฑู ประกอบด้วย สถานที่ท่องเที่ยว ที่พัก กิจกรรมท่องเที่ยว และ Tosqui-Lucks and Silva (2012) ใช้ Protégé ในการออกแบบคำศัพท์พื้นฐานที่ควรจะมีในออนโทโลจีท่องเที่ยว เช่น สถานที่ท่องเที่ยว ที่พัก การเดินทาง เป็นต้น

งานวิจัยทางด้านข้อมูลอาหาร เช่น งานวิจัยของ Fudholi, et al. (2009) ใช้ Protégé ในการออกแบบและสร้างฟัซซีออนโทโลจีสำหรับระบบแนะนำอาหารที่เหมาะสมในแต่ละมื้อแต่ละวัน โดยมีการระบุค่าที่ใช้สำหรับประมวลผล คือ เพศ อายุ ส่วนสูง น้ำหนัก เครื่องมือที่มีแอลกอฮอล์และโรคประจำตัว ซึ่งผู้ใช้สามารถให้คะแนนเมนูอาหารต่างๆ ตามออนโทโลจีเมนูอาหาร คือ รสชาติ (หวาน เค็ม เผ็ด ขม เปรี้ยว) ราคา (แพง ถูก) คะแนนโหวต (สูง ต่ำ) รวมถึงงานวิจัยของ Snae and Brueckner (2008) ใช้ Protégé ในการออกแบบออนโทโลจีอาหารใช้สำหรับเป็นความรู้ในการค้นหารายการอาหารที่เหมาะสมต่อผู้ใช้โดยยึดหลักโภชนาการอาหารและมีคลาสทั้งหมด 9 คลาส คือ Ingredients Preparation-Methods Regional-Cuisine Price Dishes Utensils Availability Nutrients และ Possibly-nutrition based-diseases เป็นต้น

สำหรับงานวิจัยที่มีการประยุกต์ใช้ Protégé ในการออกแบบและสร้างออนโทโลจีในด้านอื่นๆ เช่น งานวิจัยของ Martin and Leon (2010) ใช้ Protégé ในการออกแบบและสร้างออนโทโลจีเพื่อใช้ในห้องสมุดดิจิตอลที่เก็บรายละเอียดเช่น หนังสือ อาจารย์ บุคลากร นักศึกษา บริการการยืมคืนหนังสือและการค้นหาเชิงความหมาย ซึ่งงานวิจัยของ Wei, et al. (2010)

ใช้ Protégé ในการออกแบบและสร้างออนโทโลจีสำหรับเก็บข้อมูลในโรงพยาบาลไท้หยวน โดยมีคลาสคนที่แยกได้เป็นหมอและผู้ป่วย การรักษา และผู้ปฏิบัติงาน อีกทั้งงานวิจัยของ Chen, et al. (2011) ใช้ Protégé ในการออกแบบและสร้างออนโทโลจีเพื่อใช้ในระบบสนับสนุนการตัดสินใจสำหรับสั่งยาจากผู้แทนจำหน่ายยาเพื่อลดข้อผิดพลาดในการจัดซื้อยาและงานวิจัยของ Heeptaisong and Shivihok (2012) ใช้ Protégé ในการออกแบบและสร้างออนโทโลจีดินที่จัดเก็บองค์ความรู้เกี่ยวกับเรื่องดิน เป็นต้น

OWL (Ontology Web Language)

McGuinness and Harmelen (2004) กล่าวไว้ว่า OWL ถูกสร้างขึ้นโดย W3C Web Ontology Working Group เพื่อเป็นส่วนขยายต่อจากภาษา RDF (Resource Description Framework) และใช้ในการแก้ปัญหาข้อจำกัดในการระบุเงื่อนไขให้กับความสัมพันธ์ระหว่างคลาสในภาษา RDF (Klyne and Carroll, 2004) ซึ่ง OWL นั้นจัดเป็นองค์ประกอบหนึ่งของเว็บเชิงความหมายที่ใช้ในการบรรยายข้อมูลเชิงความหมาย สามารถกำหนดโครงสร้างข้อมูลในลักษณะลำดับชั้นและอธิบายข้อมูล (Metadata) ที่มีความสัมพันธ์ในระบบฐานข้อมูลได้ อีกทั้งยังสามารถใช้ผสมผสานกันระหว่าง RDF และ XML (eXtensible Markup Language) ขึ้นอยู่กับลักษณะของการใช้งานและรายละเอียดของข้อมูลซึ่งถูกจัดเก็บเป็นไฟล์นามสกุล .owl, .rdf และ .xml ตามลำดับ โดย Horrocks, et al. (2007) แบ่ง OWL เป็น 3 ประเภท คือ

1. OWL Lite นิยมใช้ในลักษณะลำดับชั้นของคลาสอย่างง่าย

2. OWL DL เป็นรูปแบบคำสั่งตรรกะที่สามารถตัดสินใจ สามารถจัดแยกลำดับชั้นอย่างอัตโนมัติและสามารถตรวจสอบความถูกต้องตรงกันไม่ขัดแย้งกันภายในออนโทโลจีได้

3. OWL Full เป็นภาษาที่สามารถระบุประเภทของชนิดข้อมูลว่าเป็นชนิดใดเพื่อให้รองรับการตัดสินใจหรือเป็นภาษาที่สามารถคำนวณทางตรรกะได้อย่างสมบูรณ์

สำหรับงานวิจัยนี้ได้เลือกใช้ OWL แบบ OWL Full เนื่องจากสามารถรองรับการทำงานในลักษณะการทำงานแบบตัดสินใจซึ่งอยู่ในกระบวนการอนุมานได้และใช้ออนโทโลจีท่องเที่ยวที่ได้ออกแบบ พัฒนาและจัดเก็บข้อมูลท่องเที่ยวในประเทศไทยในรูปแบบของไฟล์ OWL โดยใช้ภาษา SPARQL ในการสืบค้นข้อมูลออนโทโลจีท่องเที่ยวตามคลาสดังต่อไปนี้ คลาสสถานที่ท่องเที่ยว คลาสที่พัก คลาสร้านอาหาร คลาสร้านขายของฝากและคลาสร้านหนึ่งตำบลหนึ่งผลิตภัณฑ์ ดังแสดงตัวอย่างการเก็บข้อมูลสถานที่ท่องเที่ยว ในภาพ 2 ซึ่งจะเห็นว่าในไฟล์ OWL นี้มีการจัดเก็บข้อมูลอยู่ในรูป tag ต่างๆ เช่น tag has Name Of Attraction เก็บชื่อสถานที่ท่องเที่ยว has Other Name Of Attraction เก็บชื่อเรียกอื่นหรือชื่อเดิมของสถานที่ท่องเที่ยว has Address Of Attraction เก็บที่อยู่ของสถานที่ท่องเที่ยว has Latitude Of Attraction เก็บค่าละติจูดของสถานที่

ท่องเที่ยว has Longtitude Of Attraction เก็บค่าลองติจูดของสถานที่ท่องเที่ยว has URL Of Attraction เก็บชื่อเว็บไซต์ของสถานที่ท่องเที่ยว และ is Located In Province เก็บชื่อจังหวัดของสถานที่ท่องเที่ยว เป็นต้น ส่วนการจัดเก็บในรูปแบบภาษาไทย ภาษาอังกฤษ หรือภาษาอื่นๆ นั้น สามารถกำหนดได้ที่ property ของแต่ละ tag เช่น xml: lang = "th" หมายถึง ภาษาไทย และ xml: lang = "en" หมายถึง ภาษาอังกฤษ เป็นต้น

```
<Temple rdf:about="http://www.owl-ontologies.com/tourism.owl#วัดพระศรีรัตนมหาธาตุวรมหาวิหาร_พิษณุโลก">
   <hasNameOfAttraction xml:lang="th">วัดพระศรีรัตนมหาธาตุวรมหาวิหาร</hasNameOfAttraction>
   <hasOtherNameOfAttraction xml:lang="th">วัดใหญ่</hasOtherNameOfAttraction>
   <hasAddressOfAttraction xml:lang="th">ต.ในเมือง อ.เมือง จ.พิษณุโลก</hasAddressOfAttraction>
   <hasLatitudeOfAttraction xml:lang="th">16.823661</hasLatitudeOfAttraction>
   <hasLongtitudeOfAttraction xml:lang="th">100.262242</hasLongtitudeOfAttraction>
   <hasURLOfAttraction xml:lang="th">http://www.watyai.com/</hasURLOfAttraction>
   <hasDetailOfAttraction xml:lang="th">-</hasDetailOfAttraction>
   <isLocatedInProvince rdf:resource="http://www.owl-ontologies.com/tourism.owl#Phitsanulok"/>
</Temple>
```

ภาพ 2 แสดงตัวอย่างการเก็บข้อมูลในไฟล์ OWL (tourism.owl)

SPARQL

SPARQL อ่านว่า สปาร์เคิล เป็นภาษามาตรฐานที่ใช้สำหรับสืบค้นข้อมูลที่ถูกจัดเก็บในรูปแบบ RDF หรือ OWL (Sbodio, et al., 2010) มาแสดงหรือเรียกได้ว่าเป็นภาษาสอบถาม (Query Language) โดยอยู่บนพื้นฐานของข้อมูลที่เป็นไปในรูปแบบของกราฟที่ถูกจัดเก็บในรูปแบบของ OWL กราฟที่มีลักษณะที่ง่ายที่สุดซึ่งมีการเข้าถึงข้อมูลโดยอาศัยโครงสร้าง Triple (Basic, Graph, Pattern) สำหรับในการสืบค้นข้อมูลท่องเที่ยวนั้น จะทำได้ 2 ส่วน คือ ส่วนแสดงผลและส่วนเงื่อนไข ซึ่งในส่วนแสดงผลนั้นจะถูกแสดงไว้ที่คำสั่ง SELECT โดยตัวแปรในคำสั่ง SELECT จะมีผลต่อเงื่อนไขในคำสั่ง WHERE ดังแสดงในตาราง 2

ตาราง 2 แสดงรูปแบบและตัวอย่างภาษา SPARQL

รูปแบบภาษา SPARQL	ตัวอย่างภาษา SPARQL
SELECT ?variablelist	SELECT (COUNT(*) AS ?count)
WHERE {Basic Graph Pattern}	WHERE {?a myont:hasNameOfAttraction ?count; }

จากตาราง 2 แสดงรูปแบบและตัวอย่างภาษา SPARQL ในการนับจำนวนสถานที่ท่องเที่ยวทั้งหมดที่เก็บในออนโทโลจีท่องเที่ยว โดยใช้ฟังก์ชั่น COUNT(*) ในคำสั่ง SELECT และในส่วนของเงื่อนไขในคำสั่ง WHERE เป็นการสืบค้นข้อมูล โดยมีเงื่อนไขว่าจะนับเฉพาะชื่อสถานที่ท่องเที่ยวทั้งหมดเท่านั้นและส่งไปยังตัวแปร ?count เพื่อทำการแสดงผล

จากงานวิจัยของ Perez, et al. (2006) ที่ได้ทำการปรับปรุงรูปแบบภาษา SPARQL โดยเพิ่มเงื่อนไข AND UNION FILTER ลงในส่วนของคำสั่ง WHERE ทำให้การสืบค้นข้อมูลสามารถทำได้ซับซ้อนมากยิ่งขึ้นและเป็นการเพิ่มความสามารถในการสืบค้นของภาษา SPARQL ซึ่งในปัจจุบันมีนักวิจัยนำภาษา SPARQL มาประยุกต์ใช้สำหรับสืบค้นข้อมูลต่างๆ ได้แก่ งานวิจัยของ Quilitz and Leser (2008) ใช้ภาษา SPARQL ในการให้บริการการสืบค้นข้อมูลจากเว็บไซต์ dbpedia.org จำนวน 31.5 ล้านรายการ และงานวิจัยของ Hartig, et al. (2009) ใช้ภาษา SPARQL ในการสืบค้นข้อมูลจาก Open Data Community Project โดยทั้ง 2 งานวิจัยนี้เป็นระบบที่รองรับการสืบค้นข้อมูลผ่านระบบเครือข่ายและรองรับการเข้าถึงจากผู้ใช้หลายคนพร้อมๆ กันด้วย ต่อมา Petrak, et al. (2010) ใช้ภาษา SPARQL ในการพัฒนาเว็บสำหรับการค้นหาข้อมูลตารางปฏิบัติงาน ข้อมูลผู้แต่งหนังสือจากไฟล์ที่ถูกเก็บรูปแบบ RDF สำหรับการประยุกต์ใช้ภาษา SPARQL ในการสืบค้นข้อมูลท่องเที่ยวนั้นมีงานวิจัยของ Kongthon, et al. (2011) ใช้ภาษา SPARQL ในการสืบค้นข้อมูลท่องเที่ยวในประเทศไทย ซึ่งประกอบด้วยข้อมูลสถานที่ท่องเที่ยว ที่พัก และร้านอาหาร โดยผู้ใช้จะป้อนคำถามเพื่อสอบถามข้อมูลท่องเที่ยวตามความต้องการ เช่น "รบกวนแนะนำที่พักที่เกาะกูดหน่อยครับไปกัน 10 คน" จากนั้นระบบใช้กระบวนการตัดคำภาษาไทยโดยเทียบจากพจนานุกรมและจัดกลุ่มของคำให้เข้ารูปแบบโครงสร้างออนโทโลจีท่องเที่ยวตามที่ได้จัดเก็บไว้ แล้วสร้างชุดคำสั่งภาษา SPARQL เพื่อส่งไปสืบค้นข้อมูลออนโทโลจีท่องเที่ยวแล้วแสดงผลลัพธ์ตามที่ผู้ใช้ต้องการ

สำหรับงานวิจัยนี้ใช้ภาษา SPARQL สำหรับสืบค้นข้อมูลในออนโทโลจีท่องเที่ยวในประเทศไทยที่ได้จัดเก็บไว้ในรูปแบบไฟล์ OWL ชื่อ tourism.owl โดยพัฒนาเว็บแอพพลิเคชั่นด้วยภาษา JSP และใช้ Jena API เชื่อมต่อกับออนโทโลจีท่องเที่ยว

Jena API

Jena API เป็นเครื่องมือสำหรับนักพัฒนาเว็บเชิงความหมาย ทำหน้าที่เป็น OWL Parser ใช้งานได้แบบไม่มีค่าใช้จ่าย มีคู่มือการใช้งานพร้อมตัวอย่างโปรแกรมทำให้นักพัฒนาเว็บเชิงความหมายสามารถใช้งานได้อย่างง่าย รวมถึงยังมีความสามารถในการรองรับกลไกการอนุมาน ซึ่งผู้วิจัยใช้ Jena API เพื่ออ่านโครงสร้างข้อมูลออนโทโลจีที่ท่องเที่ยวในประเทศไทยตามที่ได้ออกแบบและจัดเก็บข้อมูลท่องเที่ยวในประเทศไทยไว้แล้วด้วยภาษา SPARQL ทั้งนี้เนื่องจาก Jena API เป็นเครื่องมือที่ถูกออกแบบให้รองรับภาษา Java และภาษา JSP (Java Server Page) ดังแสดงตัวอย่างโปรแกรมในภาพ 3

```
1. Model model = ModelFactory.createDefaultModel("C:/ontology/tourism.owl");
2. InputStream in = FileManager.get().open();  model.read(in," ");
3. String sparql = prefix + "SELECT (COUNT(*) AS ?count) " +
4. "WHERE {?a myont:hasNameOfAttraction ?count; " +
5. "myont:hasNameOfAttraction ?name; " +
6. ". FILTER (langMatches( lang(?name), '" +  lng + "') && " +
7. "(CONTAINS(str(?name),'" + keyword + "'))))}";
8. Query query = QueryFactory.create(sparql);
9. QueryExecution qe = QueryExecutionFactory.create(query,model);
10. ResultSet results = qe.execSelect();
11. if (results.hasNext()) {
12. QuerySolution row = results.next();
13. out.print( Integer.parseInt(row.getLiteral("count").getValue().toString()));
14. }
```

ภาพ 3 แสดงตัวอย่างการใช้ Jena API กับภาษา JSP

จากภาพ 3 แสดงตัวอย่างการใช้ Library ของ Jena API ซึ่งได้แก่

1. บรรทัดที่ 1 ใช้สำหรับโหลดโครงสร้างและข้อมูลออนโทโลจีท่องเที่ยว (tourism.owl)
2. บรรทัดที่ 8 ใช้สำหรับสร้างตัวแปรเพื่อเก็บรูปแบบคำสั่ง SPARQL
3. บรรทัดที่ 9 ใช้สำหรับสร้างตัวแปรเพื่อเตรียมประมวลผลคำสั่ง SPARQL
4. บรรทัดที่ 10 ใช้สำหรับเก็บผลลัพธ์จากการประมวลผลคำสั่ง SPARQL ด้วยฟังก์ชั่น execSelect( );

5. บรรทัดที่ 12 ใช้สำหรับดึงข้อมูลในออนโทโลจีมาทีละ 1 รายการ

6. บรรทัดที่ 13 ใช้สำหรับดึงข้อมูลในออนโทโลจีตามตัวแปรที่ต้องการจากรายการที่ได้ในบรรทัดที่ 12 เช่น ตัวแปร count

SWRL (Semantic Web Rule Language)

SWRL เป็นภาษาสำหรับการสร้างกฎใน Semantic Web ซึ่งเป็นภาษาที่มีการรวมคุณสมบัติของภาษา OWL และ RML ไว้ด้วยกัน ภาษา SWRL ถูกกำหนดมาตรฐานโดย W3C มีรูปแบบของกฎที่สร้างจากภาษา SWRL คือ Person(?p)^hasAge(?p,?a)^swrlb:lessThan (?a,5)->query:select(?p,?a) ซึ่งหมายความว่ากฎนี้จะแสดงคน (p) ที่มีอายุ (a) น้อยกว่า 5 ปี (O'Connor, et al., 2008) ซึ่งในงานวิจัยนี้ได้ใช้ภาษา SWRL ในการสร้างกฎต่างๆ ที่เกี่ยวข้องกับอายุ เพศ สถานที่ท่องเที่ยว จังหวัดที่ต้องการไป เพื่อประกอบการพิจารณาสำหรับแนะนำข้อมูลท่องเที่ยวในประเทศไทย โดยระบบจะให้ผู้ใช้ป้อนข้อมูลท่องเที่ยวที่สนใจ จากนั้นระบบจะประมวลผลแล้วแนะนำข้อมูลท่องเที่ยวตามความต้องการของผู้ใช้ ซึ่งผู้วิจัยใช้ Protégé API 3.5 ในการเชื่อมต่อกับ Jess Inference Engine ที่เป็นกลไกในการอนุมานเพื่อแนะนำข้อมูลท่องเที่ยวตามกฎเกณฑ์ที่ได้กำหนดไว้ในระบบแล้ว (บทที่ 3 หัวข้อการออกแบบและพัฒนาออนโทโลจีท่องเที่ยว)

หลักการออนโทโลจีเชิงเวลา (Temporal Ontology)

หลักการออนโทโลจีได้ถูกนำมาใช้กับเวลา (Time) โดยจะถูกเรียกว่า ออนโทโลจีเชิงเวลา (Temporal Ontology หรือ Ontology of Time) ส่วนใหญ่จะนำมาใช้ในการนิยามเวลาหรือใช้ในการจัดกลุ่มของข้อมูลที่มีความสัมพันธ์กัน โดยใช้เวลาเป็นเกณฑ์ในการวิเคราะห์ ซึ่งปัจจุบันมีการนำออนโทโลจีเชิงเวลาไปในงานวิจัยหลายๆ ด้าน เช่น Zhang, et al. (2007) สร้างออนโทโลจีเชิงเวลาสำหรับปฏิทินจีนที่มีทั้งหมด 9 แบบ คือ huajia, ji, lunar year, quarter, lunar month, xun, lunar day, shichen และ geng และ Zhang, et al. (2011) ได้นำเว็บเชิงความหมายมาช่วยแสดงผลปฏิทินสากลเทียบกับปฏิทินจีน ต่อมา Baratis, et al. (2009) ใช้ภาษา TOQL ที่มีความคล้ายคลึงกับภาษา SQL ในการสอบถามข้อมูลตามช่วงเวลาของเหตุการณ์ต่างๆ ที่จะเกิดขึ้น แบ่งออกเป็น 3 ช่วงเวลา คือ ช่วงเวลาในอดีต ช่วงเวลาปัจจุบันและช่วงเวลาในอนาคต และจักรกฤษณ์ เสน่ห์ นมะหุต (2557) ได้อ้างอิงงานวิจัยในการใช้ออนโทโลจีเชิงเวลาในการสร้างกฎความสัมพันธ์สำหรับการค้นหาเชื่อมโยงข้อมูลและระบุความสัมพันธ์ของความเป็นพ่อลูกเข้าด้วยกัน (Snae, 2006) เช่น คนที่จะเป็นพ่อลูกกันได้นั้นนอกจากจะมีนามสกุลเหมือนกันแล้วจะต้องมีความแตกต่างของอายุ ซึ่งออนโทโลจีเชิงเวลายังสามารถใช้ในการระบุคู่สามีภรรยาได้อีก

ด้วย โดยการใช้กฎออนโทโลจีเชิงเวลาที่ว่าผู้ที่จะเป็นสามีภรรยากันได้นั้นจะต้องมีอายุไม่น้อยกว่า 18 ปี ซึ่งจากกฎความสัมพันธ์นี้เราสามารถสรุปได้ว่าผู้ที่มีอายุต่ำกว่า 18 ปี ไม่สามารถเป็นพ่อได้

โดยงานวิจัยนี้ได้ประยุกต์ใช้ออนโทโลจีเชิงเวลาในการแนะนำข้อมูลท่องเที่ยวประเทศไทยเพื่อใช้ในการค้นหาข้อมูลท่องเที่ยวภายในจังหวัดตามฤดูกาล เช่น ภูทับเบิก ควรไปเที่ยวในช่วงฤดูหนาวเนื่องจากอากาศเย็น มีทะเลหมอกในตอนเช้าให้ชมอย่างสวยงาม และเทศกาลตามช่วงเวลา ซึ่งจะได้ไม่พลาดเทศกาลสำคัญๆ อีกทั้งแสดงข้อมูลพื้นฐานของจังหวัดนั้นๆ จะถูกแสดงอย่างมีความสัมพันธ์กันอย่างถูกต้องครบถ้วนและยังใช้ปฏิทินจันทรคติเพื่อบ่งบอกวันสำคัญต่างๆ ที่ในแต่ละปีจะไม่ตรงกัน เช่น วันวิสาขบูชา วันลอยกระทง เป็นต้น ซึ่งจะทำให้นักท่องเที่ยวไม่พลาดเทศกาล ประเพณีที่สำคัญในช่วงเวลานั้น ดังรายละเอียดต่อไปนี้

ฤดูกาลในประเทศไทย

การท่องเที่ยวแห่งประเทศไทย (2555) อธิบายถึงลักษณะภูมิอากาศของประเทศไทยว่าเป็นประเทศที่มีภูมิอากาศแบบเขตร้อน มีลมมรสุมตะวันตกเฉียงใต้ที่พัดมาจากมหาสมุทรอินเดียทำให้เกิดฤดูฝนและมีลมมรสุมตะวันออกเฉียงเหนือที่พัดจากทะเลจีนใต้ ทำให้เกิดฤดูหนาว โดยแต่ละภาคจะมีลักษณะภูมิอากาศแตกต่างกันไม่มาก ซึ่งเกิดจากปัจจัยด้านอุณหภูมิ ปริมาณน้ำฝนและความห่างไกลจากทะเล สำหรับภาคเหนือ ภาคตะวันออกเฉียงเหนือและภาคกลางจะมีลักษณะภูมิอากาศแบบสะวันนา คือ มีช่วงฤดูฝนและฤดูแล้งสลับกันชันเจน ส่วนภาคใต้และภาคตะวันออกมีภูมิอากาศแบบป่าฝนเมืองร้อน คือ ฝนตกเกือบตลอดทั้งปี ดังมีรายละเอียดการแบ่งฤดูกาลในประเทศไทยต่อไปนี้

1. ภาคเหนือ มี 3 ฤดูกาล คือ ฤดูร้อน จะอยู่ในช่วงเดือนมีนาคมถึงเมษายน ฤดูฝน จะอยู่ในช่วงเดือนพฤษภาคมถึงตุลาคม และฤดูหนาว จะอยู่ในช่วงเดือนพฤศจิกายนถึงกุมภาพันธ์

2. ภาคตะวันออกเฉียงเหนือ มี 3 ฤดูกาล คือ ฤดูร้อน จะอยู่ในช่วงเดือนกุมภาพันธ์ถึงเมษายน ฤดูฝน จะอยู่ในช่วงเดือนพฤษภาคมถึงตุลาคม และฤดูหนาว จะอยู่ในช่วงเดือนพฤศจิกายนถึงมกราคม

3. ภาคกลาง มี 3 ฤดูกาล คือ ฤดูร้อน จะอยู่ในช่วงเดือนกุมภาพันธ์ถึงเมษายน ฤดูฝน จะอยู่ในช่วงเดือนพฤษภาคมถึงตุลาคม และฤดูหนาว จะอยู่ในช่วงเดือนพฤศจิกายนถึงมกราคม

4. ภาคใต้ มีเพียง 2 ฤดูกาลเท่านั้น คือ ฤดูฝน (ฤดูมรสุม) และ ฤดูร้อน (ฤดูท่องเที่ยว) โดยฝั่งตะวันออก (ทะเลอ่าวไทย) ฤดูร้อนจะอยู่ในช่วงเดือนพฤษภาคมถึงกันยายน ส่วนฝั่งตะวันตก (ทะเลอันดามัน) ฤดูร้อนจะอยู่ในช่วงเดือนพฤศจิกายนถึงเมษายน

ปฏิทินสุริยคติและปฏิทินจันทรคติ

ปฏิทินไทยแบ่งได้เป็น 2 แบบ (ปริญญา นิ่มประยูร, 2543) คือ

1. ปฏิทินสุริยคติเป็นปฏิทินที่คำนวณขึ้นมาโดยถือเอาการโคจรของพระอาทิตย์เป็นหลัก

2. ปฏิทินจันทรคติเป็นปฏิทินที่คำนวณขึ้นมาโดยถือเอาการโคจรของพระจันทร์เป็นหลัก

ปฏิทินสุริยคติใช้หลักการโคจรของโลกรอบดวงอาทิตย์ โดยในหนึ่งรอบจะใช้เวลาประมาณ 365 1/4 วัน ถือว่าเป็นเวลา 1 ปี และการหมุนรอบตัวเองของโลกนั้นจะกินเวลาประมาณ 24 ชั่วโมง ถือว่าเป็น 1 วัน ซึ่งการหมุนรอบตัวเองของโลกทำให้เกิดมีกลางวันและกลางคืนในรอบหนึ่งๆ ขึ้นมาด้วย ในเวลา 1 ปีจะมี 12 เดือน คือ มกราคม กุมภาพันธ์ มีนาคม เมษายน พฤษภาคม มิถุนายน กรกฎาคม สิงหาคม กันยายน ตุลาคม พฤศจิกายน ธันวาคม แต่ในบางเดือนอาจมีจำนวนวันเป็น 28 หรือ 30 หรือ 31 วัน เมื่อนำจำนวนวันในแต่ละเดือนมารวมกันแล้วได้ผลลัพธ์เท่ากับ 365 วัน จะเรียกว่า "ปีปกติสุรทิน" สำหรับเดือนกุมภาพันธ์นั้นบางปีจะมี 29 วัน เมื่อนำจำนวนวันในแต่ละเดือนมารวมกันแล้วได้ผลลัพธ์เท่ากับ 366 วัน (เกิดขึ้นในทุก 4 ปี) จะเรียกว่า "ปีอธิกสุรทิน" และสำหรับวิธีการคำนวณหาปีอธิกสุรทินทำได้โดยนำตัวเลข 2 ตัวท้ายของปี ค.ศ. แล้วนำมาหารด้วย 4 หากหารแล้วลงตัวให้ถือว่าปีนั้นเป็น "ปีอธิกสุรทิน" แต่กรณีที่เป็นปี พ.ศ. ให้ทำการลบด้วย 543 เพื่อแปลงเป็นปี ค.ศ. ก่อนทำการคำนวณ

ปฏิทินจันทรคติกำหนดขึ้นมาโดยใช้หลักการโคจรของพระจันทร์ที่โคจรไปรอบโลกมาเป็นหลัก ซึ่งการโคจรของพระจันทร์ที่หมุนไปรอบโลกนั้น มีข้อสังเกตได้ 2 ข้อ คือ

1. พระจันทร์โคจรเข้าสู่จุดเดิม ใช้เวลาประมาณ 22 วันเศษ

2. พระจันทร์โคจรเข้าร่วมราศีกับอาทิตย์อีกจุดหนึ่ง ใช้เวลาประมาณ 29 วันกับ 12 ชั่วโมง 44 นาที 2.9 วินาที

ในการกำหนดเดือนทางจันทรคตินั้นถือเอาการโคจรของพระจันทร์ในข้อที่ 2 โดยในแต่ละเดือนจะมี 29 วันหรือ 30 วันสลับกันไป ดังรายละเอียดต่อไปนี้

1. เดือน "คี่" คือเดือน 1 (อ้าย), 3, 5, 7, 9, 11 ให้มี 29 วัน และเรียกว่า "เดือนขาด"

2. เดือน "คู่" คือเดือน 2 (ยี่), 4,6,8,10,12 ให้มี 30 วัน และเรียกว่า "เดือนเต็ม"

สำหรับการนับวันทางจันทรคติเรียกว่า "ดิถี" โดยแบ่งเป็น 2 ภาค คือ

1. ภาคแรกให้มี 15 วัน เรียกว่า ข้างขึ้น

2. ภาคหลังให้มี 14 วัน (เดือนขาด) และให้มี 15 วัน (เดือนเต็ม) เรียกว่า ข้างแรม

จากการเรียกวันทางจันทรคติที่แบ่งเป็นภาคแล้ว เรียกว่า ค่ำ ตามภาค ซึ่งวันและเดือนทางจันทรคติบางครั้งไม่เป็นไปตามกฎ เช่น เดือน 7 บางเดือนมี 29 วัน แต่ในปี พ.ศ. 2513 เดือน 7 กลับมี 30 วัน เรียกว่า ปีอธิกวาร และในบางปีจำนวนเดือนในรอบปีเป็น 13 เดือน คือ มีเดือน 8

ซ้ำกัน 2 หน เช่นในปี 2512, 2535 เป็นต้น ทั้งนี้เกิดขึ้นมาจากวันในรอบปีทางจันทรคติที่น้อยกว่าวันในรอบปีของสุริยคติประมาณปีละ 10 วัน ดังนั้นในเวลาประมาณ 3 ปี วันในปีทางจันทรคติจะน้อยกว่าวันในทางสุริยคติ 30 วัน ทำให้มีการเพิ่มเดือนทางจันทรคติให้ซ้ำกันขึ้นมาเดือนหนึ่งเพื่อให้จำนวนวันทางจันทรคติและสุริยคติเท่ากัน โดยกำหนดให้เป็นเดือน 8 และเดือน 8 ที่ถูกเพิ่มขึ้นมาจะเรียกว่า แปดหลัง แล้วเรียกปีนั้นว่า ปีอธิกมาส จากกรณีตัวอย่างข้างต้นนี้สามารถสรุปวันทางจันทรคติได้ 3 ลักษณะ คือ

1. ปีปกติมาส ปกติวาร เป็นปีที่มี 12 เดือน ประกอบด้วยเดือนเต็ม 6 เดือนและเดือนขาด 6 เดือน รวมทั้งปีมี 354 วัน

2. ปีปกติมาส อธิกวาร เป็นปีที่มี 12 เดือน ประกอบด้วยเดือนเต็ม 7 เดือน โดยในเดือนที่ 7 จะมี 30 วันและเดือนขาด 5 เดือน รวมทั้งปีมี 355 วัน

3. ปีอธิกมาส ปกติวาร เป็นปีที่มี 13 เดือน โดยมีเดือน 8 สองเดือน (แปดหน้าแปดหลัง) ทำให้ปีทั้งปีมี 384 วัน

หลักเกณฑ์ในการคำนวณปฏิทินจันทรคติในปฏิทินโหราศาสตร์ได้ยึดหลักคำนวณจากวันออกพรรษาเป็นหลัก เนื่องจากคือวันปวารณาออกพรรษาจะต้องเป็นวันเพ็ญและตามปฏิทินจันทรคติของไทย ซึ่งวันออกพรรษานั้นจะตรงกับวันขึ้น 15 ค่ำเดือน 11 ดังนั้นเมื่อรู้วันที่ออกพรรษาในแต่ละปีแล้ว จะทำให้รู้ว่าปีนั้นเป็นปีอะไร เช่น ปกติมาส หรือ ปีอธิกวาร หรือ ปีอธิกมาส โดยในการคำนวณนั้นแสดงดังสมการที่ 1 (จำรัส ศิริ, 2549)

$$days=\sum_{i=1}^{12} dm_i + d1 - d2 \qquad (1)$$

เมื่อ days คือ จำนวนวันใน 1 ปีของปีที่คำนวณ

d1 คือ วันที่ออกพรรษาในปีที่คำนวณ

d2 คือ วันที่ออกพรรษาของปีที่แล้ว

ตัวอย่างเช่น ในปี พ.ศ. 2556 วันออกพรรษานั้นที่ตรงกับ 15 ค่ำ เดือน 11 คือวันที่ 19 ตุลาคม และปี พ.ศ. 2557 จะตรงกับวันที่ 8 ตุลาคม จากการคำนวณโดยใช้สมการที่ 1 จะได้ผลลัพธ์ดังนี้ 365 + 8 - 19 = 354 วัน สรุปได้ว่าปี 2557 นี้เป็นปีปกติมาส ปกติวาร

จากนั้นนำปีที่ได้นี้ไปทำการคำนวณทุกๆ ปีล่วงหน้าเพื่อจะทราบวันสำคัญทางศาสนาตรงกับวันที่เท่าไหร่นั้นของแต่ละปี เพื่อใช้สำหรับการแนะนำข้อมูลท่องเที่ยวตามเทศกาลท่องเที่ยวและประเพณีที่มีในประเทศไทย โดยในงานวิจัยนี้ได้พัฒนาอัลกอริทึมสำหรับระบบแนะนำข้อมูลท่องเที่ยวในประเทศไทยตามช่วงเวลา ซึ่งใช้ปฏิทินจันทรคติในการแนะนำเทศกาลหรือประเพณีสำคัญๆ ประจำจังหวัด ซึ่งวันดังกล่าวนั้นในแต่ละปีจะไม่ตรงกัน เช่น ขึ้น 15 ค่ำ เดือน 12 เป็นวันลอยกระทง ในปี พ.ศ. 2556 ตรงกับวันที่ 17 พฤศจิกายน ในปี พ.ศ. 2557 ตรงกับวันที่ 6 พฤศจิกายน และปี 2558 ตรงกับวันที่ 25 พฤศจิกายน โดยผู้วิจัยใช้สมการ 1 ในการคำนวณหาจำนวนวันในแต่ละปีเพื่อให้ทราบลักษณะวันทางจันทรคติ จากนั้นนำไปใช้ในการพัฒนาอัลกอริทึมสำหรับคำนวณปฏิทินจันทรคติ (ดูเพิ่มเติมในบทที่ 3 หัวข้อการแนะนำข้อมูลท่องเที่ยวตามช่วงเวลาและบริเวณใกล้เคียง)

ทฤษฎีเนมแมทชิ่ง (Name Matching)

Snae (2006) กล่าวว่าการเปรียบเทียบระหว่างชื่อ 2 ชื่อเพื่อให้รู้ว่าคล้ายกัน เหมือนกันหรือแตกต่างกันหรือไม่นั้นอาจใช้วิธีการคำนวณค่าความเหมือน ค่าความคล้าย ค่าความแตกต่าง ซึ่งสามารถใช้ตัดสินใจและเป็นข้อสรุปของชื่อ 2 ชื่อจะเหมือนกัน คล้ายกัน หรือแตกต่างกันหรือไม่ การใช้เนมแมทชิ่งในการเปรียบเทียบชื่อ 2 ชื่ออาจทำได้ไม่ง่ายนัก บางทีอาจจะใช้ การออกเสียงของชื่อเป็นตัวกำหนดความเปรียบเทียบค่าความคล้ายคลึง (Similarity) หรืออาจจะใช้การเขียนและการสะกดคำเป็นตัวกำหนด หรือบางทีก็อาจจะใช้ค่าระยะห่างและเมทริกในการคำนวณหาค่าความแตกต่างซึ่งทั้งนี้ทั้งนั้นอาจจะใช้หลักการทั้งหมดมารวมกันก็ได้ โดย Snae and Brueckner (2006) ได้ทำการแบ่งชนิดของเนมแมทชิ่งไว้เป็น 4 ประเภทอย่างเห็นได้ชัดคือ กฎพื้นฐานของตัวสะกด (Spelling Base Algorithm) กฎพื้นฐานของการออกเสียง (Phonetic Base Algorithm) วิธีการรวมกัน (Composite (C) Methods) และวิธีการผสม (Hybrid Approaches)

กฎพื้นฐานของตัวสะกด (Spelling Base Algorithm)

Snae (2006) กล่าวว่าอัลกอริทึมนี้ใช้การเปรียบเทียบตัวอักษรระหว่างชื่อ 2 ชื่อ ว่ามีหลักการเขียนคล้ายกันอย่างไร อัลกอริทึมประเภทนี้ได้แก่ กัส (Guth) และเลเวลสไตล์ (Levanshtein) และผลของการเปรียบเทียบนี้จะออกมาในรูปว่าคล้ายคลึงหรือเหมือนกันหรือไม่ (Match or Nonmatch)

Levenshtein Distance (LD) หรือเรียกว่า ค่าระยะห่างของการเปลี่ยน (Edit Distance) เป็นการวัดค่าความแตกต่างจำนวนการกระทำระหว่างชื่อ 2 ชื่อ (Levenshtein, 1965) ว่ามีการกระทำกี่ครั้ง ซึ่งจำนวนครั้ง คือค่าระยะห่างของการเปลี่ยน (Edit Distance) จะเห็นได้ว่าใน 2 ชื่อนี้

มีความแตกต่างกันอยู่ 4 ดังนั้นถ้าเรากำหนดว่าชื่อ 2 ชื่อนี้มีความแตกต่างน้อยกว่า 3 ก็แสดงว่าชื่อนี้ไม่คล้ายคลึงกัน (No Match) เพราะชื่อนี้มีความแตกต่างกันเท่ากับ 4 โดยที่ Levenshtein Algorithm นั้นจะใช้กระบวนการเพิ่ม การลด และการแทนที่ของตัวอักษรจากชื่อหนึ่งไปยังอีกชื่อหนึ่ง ตัวอย่างเช่น (Snae, 2006)

S1 มีค่าเป็น CHAKKRIT

S2 มีค่าเป็น CHAK-RAVI

Cost of LD = 4

จาก S1 และ S2

แมท (Match) : (C,C)

แมท (Match) : (H,H)

แมท (Match) : (A,A)

แมท (Match) : (K,K)

แทรก (Insert) : (-,K)

แมท (Match) : (R,R)

แทนที่ (Replace) : (A,I)

แทนที่ (Replace) : (V,T)

ลบ (Delete) : (I,-)

กฎพื้นฐานของการออกเสียง (Phonetic Base Algorithm)

Snae (2006) กล่าวว่าอัลกอริทึมนี้ใช้โครงสร้างของเสียงเป็นตัวเปรียบเทียบค่าความคล้ายกัน โดยชื่อ 2 ชื่ออาจเขียนไม่เหมือนกันแต่มีการออกเสียงเหมือนกัน อัลกอริทึมประเภทนี้ได้แก่ Soundex Metaphone NYSIIS และ Phonex อัลกอริทึมเหล่านี้สามารถแก้ปัญหาชื่อที่มีเสียงเดียวกันแต่เขียนหรือสะกดต่างกัน โดยใช้วิธีการจัดกลุ่มตัวอักษรจัดให้เสียงเดียวกันอยู่ในกลุ่มตัวอักษรเดียวกัน ตัวอย่างเช่น ตัวอักษร Ph ให้จัดอยู่ในกลุ่มตัวอักษร F หรือในภาษาไทย "จ ท ธ ต ฏ" ให้จัดอยู่ในกลุ่มแม่กด (ด) อาทิเช่น คำว่า กาจ กับ กาด จะออกเสียงเหมือนกัน

ซาวด์เด็ค (Soundex Algorithm)

Snae (2006) กล่าวว่า Soundex หรือ Russell Soundex Code (Odell and Russell, 1922) เป็นการออกแบบแรกเริ่มสำหรับใช้กับชื่อภาษาอังกฤษ โดยทำการจัดกลุ่มเสียงของคำ ซึ่งกลุ่มคำที่มีเสียงตัวสะกดเหมือนกันจะถูกจัดอยู่ในกลุ่มคำเดียวกัน (Winchester, 1970) การทำงานของอัลกอริทึมนี้จะทำการเปลี่ยนชื่อเป็น 4 ตัวอักษรของ Soundex ซึ่งประกอบด้วย 1 ตัวอักษรและ 3 ตัวเลข โดยมีกฎดังต่อไปนี้ (Snae, 2006)

1. เก็บตัวอักษรตัวแรกของ Input ชื่อไว้
2. ถ้าตัวอักษร A, E, I, O, U ไม่ใช้ตัวอักษรตัวแรกให้ตัดออกไป
3. จากข้อที่ 1 เปลี่ยนอักษรที่เหลือเป็นรหัส Soundex ดังต่อไปนี้

Letter:	Number is coded to:
B, P, F and V	1
C, G, J, K, Q, S, X and Z	2
D and T	3
L	4
M and N	5
R	6

4. ผลลัพธ์ของรหัส Soundex ต้องอยู่ในรูปของ ตัวอักษร ตัวเลข
5. จากข้อ 4 ถ้าผลลัพธ์ของรหัส Soundex น้อยกว่า 4 ตัวรหัส Soundex ให้ใส่ศูนย์แทนจบครบรหัส แต่ถ้ารหัส Soundex มากกว่า 4 รหัสให้ตัดออก

ตาราง 3 แสดงตัวอย่างของชื่อจังหวัดกับรหัส Soundex

ชื่อจังหวัด	รหัส Soundex
Lampang	L515
Maehongson	M525
Phichit	P230
Sisaket	S223

วิธีการรวมกัน (Composite Methods)

Snae (2006) กล่าวว่าอัลกอริทึมนี้เป็นการรวมกันระหว่างกฎพื้นฐานของตัวสะกดและกฎพื้นฐานของการออกเสียง ซึ่งจะทำการเปรียบเทียบตัวอักษรระหว่างชื่อ 2 ชื่อนั้นมีหลักการเขียนคล้ายกันอย่างไรกับการใช้โครงสร้างของเสียงเป็นตัวเปรียบเทียบค่าความคล้าย โดยอัลกอริทึมนี้ได้แก้ปัญหาชื่อที่เขียนเหมือนกันแต่ออกเสียงต่างกันและชื่อที่เขียนต่างกันแต่ออกเสียงเหมือนกัน อัลกอริทึมประเภทนี้ได้แก่ SIMPLEX และ ISG (Index of Similarity Group)

ISG เป็นวิธีที่จะใช้ในการวัดหรือเปรียบค่าความคล้ายคลึงกันระหว่างชื่อโดยมีสเกลจาก 0-1 ซึ่งการวัดค่าความคล้ายคลึงนั้นจะใช้สูตรในการคำนวณดังต่อไปนี้

$$\frac{I}{I+D} \qquad (2)$$

เมื่อ I คือ จำนวนตัวอักษรที่เหมือนกัน

D คือ ตัวอักษรที่แตกต่างกันระหว่างชื่อ

ตาราง 4 แสดงตัวอย่างการคำนวณชื่อจังหวัดกับ ISG

ชื่อจังหวัด	ค่า I.S.G.
Phitsanulok	1.0000
Pitsanulok	0.9091
Phitsanoolok	0.8333
Phitsanoklok	0.8333
Phisanooklok	0.7500
พิษณุโลก	1.0000
พิษนุโลก	0.8750
พิดสะนุโลก	0.6000
พิดษะณุโลก	0.8000
พิษณุโรค	0.5000

จากงานวิจัยของ กนกกาญจน์ นมะหุต (2550) ที่ใช้อัลกอริทึม ISG ในการตรวจสอบชื่อไทยในระบบการเปลี่ยนชื่อให้สอดคล้องกับชื่อพ่อและชื่อแม่ โดยนำตัวอักษรจากชื่อพ่อและชื่อแม่มาทำการแยกตัวอักษรแล้วนำอักษรที่ได้ไปเปรียบเทียบชื่อจากฐานข้อมูลแล้วคิดค่าความน่าจะเป็นว่าชื่อที่ได้มาใกล้เคียงหรือเหมือนตัวอักษรของชื่อพ่อและชื่อแม่

วิธีการผสม (Hybrid Approaches)

Snae (2006) กล่าวว่าอัลกอริทึมนี้เป็นการนำอัลกอริทึมหลายๆ แบบมารวมกัน อัลกอริทึมประเภทนี้ได้แก่ LIG (Levenshtein Index of Similarity and Guth) ซึ่งถูกพัฒนาโดย จักรกฤษณ์ (Snae and Diaz, 2002) โดยใช้ค่าความน่าจะเป็นในการวัดค่าความคล้ายกันของชื่อ ทำให้ค่าความถูกต้องในการหาชื่อคล้ายกันหรือเหมือนกันมีประสิทธิภาพมากขึ้น ในปัจจุบันได้มี พัฒนาเป็น LIG3 และสามารถใช้ได้ดีกับชื่อหลายประเทศ

ข้อดีของอัลกอริทึมนี้คือทำให้ค่าความถูกต้องในการหาชื่อคล้ายหรือชื่อที่เหมือนมีมาก ขึ้น สามารถใช้ได้ดีกับชื่อหลายๆ ประเทศ และใช้หลักการความน่าจะเป็นในการวัดค่าความ คล้ายคลึงของชื่อจะทำให้ง่ายในการออกแบบ แต่ จักรกฤษณ์ เสน่ห์ และกนกกาญจน์ นมะหุต (2550) พบว่าอัลกอริทึม LIG3 เป็นอัลกอริทึมที่ยุ่งยากและซับซ้อนในการหาชื่อคล้ายกัน เพราะว่า ประกอบด้วยหลายฟังก์ชั่น คือ ฟังก์ชั่นการคำนวณหาค่าระยะทางที่แตกต่างระหว่างชื่อสองชื่อ (หรือค่า C) ด้วยหลักการของ Levenshtein (1965) ซึ่งอัลกอริทึม LIG3 มีสูตรการคำนวณ ดังต่อไปนี้ (Snae and Diaz, 2002)

$$\frac{2l}{2l+C} \qquad (3)$$

เมื่อ l คือ จำนวนตัวอักษรที่เหมือนกันระหว่างชื่อ

C คือ ฟังก์ชั่นการคำนวณหาค่าระยะทางที่แตกต่างระหว่างชื่อ 2 ชื่อ โดยใช้ หลักการของ Levenshtein (1965) และฟังก์ชั่นการคำนวณความคล้ายคลึงของชื่อ

ตาราง 5 แสดงตัวอย่างการคำนวณชื่อจังหวัดของ LIG3

ชื่อจังหวัด	ค่า LIG
Phitsanulok	1.0000
Pitsanulok	0.9524
Phitsanoolok	0.9091
Phitsanoklok	0.9091
Phisanooklok	0.8182
พิษณุโลก	1.0000
พิษนุโลก	0.9333
พิดสะนุโลก	0.7500
พิดษะณุโลก	0.8889
พิษณุโรค	0.8571

โดยงานวิจัยนี้ได้ใช้วิธีการรวมกันจากกฎพื้นฐานของตัวสะกดและกฎพื้นฐานของการออกเสียงในการสร้างคำที่สามารถเขียนได้หลากหลายทั้งภาษาไทยและภาษาอังกฤษ ซึ่งพัฒนาเป็นอัลกอริทึมใหม่เพื่อแก้ปัญหาการพิมพ์ชื่อผิดหรือออกเสียงผิดหรือมีการเขียนไม่ถูกต้องตามหลักภาษาไทยและภาษาอังกฤษ (Name Variation Matching Algorithm) โดยมีหลักการคือเมื่อนักท่องเที่ยวป้อนคำค้นหา ระบบจะทำการสร้างคำต่างๆ ที่สามารถเขียนได้หลายรูปแบบ เพื่อจะนำไปแก้ปัญหาเว็บไซต์ที่มีการเขียนหรือสะกดผิดและไม่สามารถสืบค้นเจอให้ได้ผลลัพธ์ที่ไม่ผิดพลาดตรงกับความต้องการของผู้ใช้และใช้อัลกอริทึม ISG มาใช้เพื่อหาค่าความคล้ายคลึงของคำค้นหาที่มีการสะกดผิดหรือเขียนผิดซึ่งปรากฏอยู่ในเว็บไซต์ท่องเที่ยวต่างๆ บนอินเทอร์เน็ตและกรองคำค้นหาที่ถูกสร้างจาก Name Variation Matching เนื่องจากเป็นอัลกอริทึมที่ไม่ซับซ้อน รวมถึงจากงานวิจัยของ เกรียงกมล คำมา และจักรกฤษณ์ เสน่ห์ นมะหุต (2555) ที่มีการตั้งค่าเขตแดนไว้ที่ 0.5 ในการวัดค่าความอนาจารของคำ โดยค่าที่คำนวณนั้นมากกว่าหรือเท่ากับค่าเขตแดน 0.5 แสดงว่าชื่อมีความคล้ายคลึงกัน แต่ถ้าน้อยกว่า 0.5 แสดงว่าชื่อดังกล่าวไม่มีความคล้ายคลึงกัน โดยงานวิจัยนี้ได้ใช้ค่า ISG ที่มีค่ามากกว่าหรือเท่ากับ 0.5 มากรองชื่อที่มีความคล้ายคลึงกันก่อนส่งไปสืบค้นในเว็บไซต์กูเกิ้ล ทั้งนี้เพื่อลดปริมาณของข้อมูลที่อาจไม่เกี่ยวข้อง

อัลกอริทึมในการจัดหมวดหมู่และแนะนำข้อมูลท่องเที่ยว (Classification Algorithm and Tourism Recommendation)

ในส่วนนี้จะกล่าวถึงอัลกอริทึมที่ใช้ในการจัดหมวดหมู่เว็บไซต์ท่องเที่ยวและอัลกอริทึมในการแนะนำข้อมูลท่องเที่ยว ดังรายละเอียดต่อไปนี้

1. อัลกอริทึมในการจัดหมวดหมู่เว็บไซต์ท่องเที่ยว

การวิเคราะห์เว็บไซต์นั้น (Gore and Pitale, 2013) ได้อธิบายหลักการวิเคราะห์เว็บไซต์ไว้ 3 วิธี คือ

1.1 วิเคราะห์เนื้อหาเว็บไซต์ ได้แก่ ข้อความ รูปภาพ เสียง วิดีโอ ที่ปรากฏอยู่ในหน้าเว็บเพจ

1.2 วิเคราะห์การใช้เว็บ ได้แก่ ข้อมูลที่ถูกจัดเก็บในเครื่องของผู้ใช้ เช่น Logs file, Cookies, user profile เป็นต้น

1.3 วิเคราะห์โครงสร้างเว็บ ได้แก่ ลิงค์ที่ปรากฏอยู่ภายในหน้าเว็บเพจ ลิงค์ที่มีการเชื่อมโยงไปยังเว็บอื่น

เกรียงกมล คำมา และจักรกฤษณ์ เสน่ห์ นมะหุต (2555) ได้ใช้การวิเคราะห์เนื้อหาเว็บไซต์และการวิเคราะห์โครงสร้างเว็บมาใช้ในการพัฒนาระบบวิเคราะห์เว็บไซต์อนาจารด้วยกลุ่มคำเชิงความหมายคำอนาจารใน HTML Tags เพื่อป้องกันการเข้าถึงเว็บไซต์อนาจารจากบุคคลในองค์กรหรือกลุ่มครอบครัวที่มีเยาวชนอยู่ในบ้าน โดยใช้วิธีการคำนวณหาความถี่ของคำอนาจารภาษาไทยที่ปรากฏใน HTML Tags ในส่วนของหัวเว็บเพจ ตัวเว็บเพจและการเชื่อมโยงไปยังเว็บอื่น เฉพาะหน้าเริ่มต้นของเว็บไซต์เท่านั้น จากนั้นนำมาคำนวณหาค่าเรทติ้งด้วยเทคนิคทางสถิติเพื่อวิเคราะห์ว่าเป็นเว็บอนาจารหรือไม่ แต่ยังไม่สามารถวิเคราะห์เว็บไซต์ที่มีเนื้อหาบางส่วนหรือทั้งหมดเป็นภาษาอังกฤษได้ จึงทำให้ เกรียงกมล คำมา และจักรกฤษณ์ เสน่ห์ นมะหุต (2556) พัฒนาระบบคัดกรองเว็บไซต์อนาจารด้วยเทคนิคการวิเคราะห์เว็บไซต์เพื่อแก้ปัญหาดังกล่าว พบว่ามีประสิทธิภาพสูงถึง 94% อย่างไรก็ตามเนื่องจากงานวิจัยทั้ง 2 นี้มีการนำส่วนการเชื่อมโยงไปยังเว็บไซต์อื่นมาวิเคราะห์ทำให้เกิดการทำงานที่ล่าช้า

สำหรับการวิเคราะห์เว็บไซต์เพื่อจัดหมวดหมู่เว็บไซต์ท่องเที่ยวนั้น ผู้วิจัยใช้อัลกอริทึมในการจัดหมวดหมู่เว็บไซต์ท่องเที่ยวนั้นด้วยอัลกอริทึม LSI อัลกอริทึมนาอีฟเบย์และอัลกอริทึมนาอีฟเบย์ที่ปรับปรุงโดยใช้ค่าขอบเขต ดังรายละเอียดต่อไปนี้

Latent Semantic Indexing (LSI)

Latent Semantic Indexing (LSI) เป็นอัลกอริทึมหนึ่งในการประมวลผลภาษาธรรมชาติที่มีลักษณะเฉพาะในการหาความหมายที่แอบแฝงด้วยเหตุนี้จึงมีการนำมาใช้กับ

การสืบค้นสารสนเทศ เนื่องจากในยุคแรกนั้นใช้คำสำคัญในการระบุสิ่งที่ต้องการค้นหา แต่ปัญหาที่พบคือ บางเอกสารไม่มีคำที่ผู้ใช้ระบุแต่เอกสารนั้นตรงกับความต้องการจะไม่ถูกค้นพบ LSI นั้นเป็นวิธีการที่มีหลักการทำงานอยู่บนพื้นฐานของการคำนวณทางสถิติ โดยพิจารณาจากการปรากฏร่วมของคำต่างๆ ในการทำดัชนี (Thongkrau and Lalitrojwong, 2010) ซึ่งมีนักวิจัยนำอัลกอริทึม LSI ไปประยุกต์ใช้ในงานต่างๆ ได้แก่ Zelikovitz and Kogan (2006) ใช้อัลกอริทึม LSI ในการจัดหมวดหมู่เว็บไซต์เกี่ยวกับสัตว์ ได้แก่ นก สุนัข แมว ปลาและสัตว์ป่า เพื่อแก้ปัญหาการกระจายของเว็บไซต์เกี่ยวกับสัตว์ให้ถูกจัดอย่างเป็นหมวดหมู่ ทำให้สะดวกต่อการเข้าถึงเว็บไซต์ดังกล่าวและมีประโยชน์ต่อผู้ที่สนใจเรื่องนี้เป็นอย่างมาก ต่อมา Lucia, et al. (2007) ใช้อัลกอริทึม LSI หาความคล้ายคลึงของเนื้อหาเว็บเพจเพื่อจัดกลุ่มเนื้อหาของเว็บไซต์ที่มีความคล้ายคลึงกันให้อยู่ด้วยกัน จากกลุ่มตัวอย่าง 4 เว็บไซต์ คือ การประชุมวิชาการระดับนานาชาติด้านวิศวกรรมซอฟต์แวร์และวิศวกรรมความรู้ครั้งที่ 14 จำนวน 157 เพจ เครือข่ายโรงเรียนนานาชาติด้านวิศวกรรมซอฟต์แวร์ จำนวน 85 เพจ ข้อมูลนักศึกษาด้านวิทยาศาสตร์การเมืองมหาวิทยาลัยซาเลร์โน 449 เพจ เว็บไซต์ EasyClinic จำนวน 161 เพจ ซึ่งพบว่ามีความเหมาะสมในการหาค่าความคล้ายคลึงระหว่างเอกสารเนื่องจากง่าย สะดวกในการนำไปประยุกต์ใช้ในระบบค้นหาและเหมาะกับข้อมูลที่มีจำนวนมาก จากนั้น Qi and Davison (2007) ใช้อัลกอริทึม LSI จัดหมวดหมู่เว็บไซต์ทางด้านธุรกิจ กีฬา ศิลปะและคอมพิวเตอร์ เพื่อแก้ปัญหาการสืบค้นข้อมูลให้มีประสิทธิภาพมากขึ้น ซึ่งถือได้ว่าการใช้อัลกอริทึมนี้สามารถลดขนาดของเอกสารที่ได้รับความนิยมทำให้เอกสารมีขนาดเล็ก ใช้พื้นที่น้อย และ Zhang, et al. (2008) ใช้อัลกอริทึม LSI ในการจัดกลุ่มข้อความในภาษาจีนและภาษาอังกฤษ ซึ่งในการสร้างเมทริกซ์ของคำได้คูณค่าน้ำหนักของคำด้วยวิธี TF*IDF (วิชุดา โชติรัตน์ และคณะ, 2554) โดยใช้เอกสารที่เป็นภาษาอังกฤษจากเว็บไซต์ http://www.research.att.com/~lewis แล้วเลือกมา 4 หมวดหมู่ ได้แก่ เกษตร การค้า น่าสนใจและทั่วไป จำนวน 2,042 เอกสารและภาษาจีนจากเว็บไซต์ http://www.searchforum.org.cn/tansongbo/corpus.htm ได้เลือกมา 4 หมวดหมู่ได้แก่ เกษตร ประวัติศาสตร์ การเมืองและเศรษฐกิจ จำนวน 1,200 เอกสาร จากผลการทดลองพบว่าการใช้อัลกอริทึม LSI มีความถูกต้องเฉลี่ยในการจัดกลุ่มข้อความภาษาอังกฤษดีกว่าภาษาจีน คือ 61.25% และ 53.74% ตามลำดับ อีกทั้ง เรือนทิพย์ ทองใบ และกาญจนา วิริยะพันธ์ (2556) ได้ประยุกต์ใช้อัลกอริทึม LSI และ Vector Space Model (VSM) ในการหาความคล้ายคลึงของคำตอบที่มีในฐานข้อมูลและคำถามของผู้ใช้ โดยการพัฒนาระบบถามตอบด้านการบริการนักศึกษาเพื่อให้เกิดความสะดวกแก่ผู้ใช้ด้วยการรับคำขอและได้ผลลัพธ์ตรงตามความต้องการ จากการเปรียบเทียบประสิทธิภาพพบว่าค่า F-Measure ของอัลกอริทึม LSI (95%) มีประสิทธิภาพ

ดีกว่า VSM (77%) เนื่องจากสามารถใช้คำที่เป็นคำสำคัญหรือคำที่มีความหมายในเชิงเนื้อหาได้เป็นอย่างดีทำให้ได้ข้อมูลที่ตรงกับความต้องการของผู้ใช้มากกว่า ซึ่งสอดคล้องกับงานวิจัยของ Bhat, et al. (2004) ที่กล่าวว่าปัญหาเรื่องชื่อที่พบจากการรวบรวมข้อมูลจากเว็บไซต์มักมีการเขียนชื่อที่แตกต่างกันไป แต่หมายถึงชื่อเดียวกันหรือสถานที่เดียวกัน เช่น เมืองในประเทศอินเดีย "Bombay" บางครั้งก็พบว่าใช้ชื่อ "Mumbai" และ "บ้าน" กับ "เรือนไทย" หรือ "Human" กับ "User" มีความสัมพันธ์กันในเชิงความหมายไปในทิศทางเดียวกัน (กนกรัตน์ จิรสัจจานุกูล และณัฐธนนท์ หงส์วริทธิ์ธร, 2555)

อัลกอริทึมนาอีฟเบย์

อัลกอริทึมนาอีฟเบย์ใช้ความน่าจะเป็นในการทำนายผลโดยเป็นอัลกอริทึมที่ใช้ในการแก้ปัญหาแบบการจำแนกประเภทที่สามารถคาดการณ์ผลลัพธ์ได้ กระบวนการทำงานของอัลกอริทึมไม่มีความซับซ้อน (นฤพนธ์ พนาวงศ์ และจักรกฤษณ์ เสน่ห์ นมะหุต, 2556ค) ซึ่งอัลกอริทึมนี้แสดงดังสมการต่อไปนี้ (Patil and Pawar, 2012)

$$Cmap = \underset{c \in C}{\operatorname{argmax}} \left(P(c) \prod_{1 \le k \le n} P(t_k|c) \right) \tag{4}$$

เมื่อ Cmap คือ ค่าสูงสุดของผลคูณความน่าจะเป็นระหว่าง P(c) กับ $P(t_k \mid c)$

C คือ หมวดหมู่ทั้งหมดที่ต้องการจัด $(c_1,c_2,c_3,c_4,c_5,c_6)$

P(c) คือ ค่าความน่าจะเป็นในแต่ละหมวดหมู่ คำนวณได้จากสมการ 5

$P(t_k \mid c)$ คือ ค่าความน่าจะเป็นของความถี่ของคำที่ k ซึ่งปรากฏในเว็บไซต์ของหมวดหมู่ c โดยกำหนดให้ c มี 6 หมวดหมู่ คือ สถานที่ท่องเที่ยว ที่พัก ร้านอาหาร ร้านขายของฝาก ร้านสินค้าหนึ่งตำบลหนึ่งผลิตภัณฑ์และเทศกาล ซึ่งสามารถคำนวณได้จากสมการ 6

n คือ จำนวนคำที่ใช้ในการจัดหมวดหมู่

$$P(c) = \frac{N_c}{N} \tag{5}$$

เมื่อ N_c คือ จำนวนเว็บไซต์เรียนรู้ที่อยู่ในหมวดหมู่ c

N คือ จำนวนเว็บไซต์เรียนรู้ทั้งหมด

$$P(t_k|c)=\frac{t_k+1}{\sum_{k=1}^{n} t_k+1} \quad (6)$$

แต่เมื่อคำนวณหาความน่าจะเป็นจากสมการที่ 4 แล้วจะเกิดปัญหาทศนิยมจำนวนมาก จึงได้ปรับเป็นฟังก์ชั่น log ดังสมการ 7 (Patil and Pawar, 2012)

$$Cmap=\underset{c \in C}{\operatorname{argmax}}\left(P(c) \sum_{k=1}^{n} \log(P(t_k|c)) \right) \quad (7)$$

โดยที่ $P(c) \sum_{k=1}^{n} \log(P(t_k|c))$ เรียกว่า ค่าความน่าจะเป็นของผลคูณความน่าจะเป็นระหว่าง P(c) กับ $P(t_k \mid c)$ หรือ เรียกสั้นๆ ว่า ค่าความน่าจะเป็นของนาอีฟเบย์ (P(NB))

2. อัลกอริทึมสำหรับแนะนำข้อมูลท่องเที่ยว

อัลกอริทึม K-Nearest Neighbor (K-NN) เป็นการจำแนกหรือจัดกลุ่มโดยพิจารณาจากชุดข้อมูลใกล้เคียง กล่าวคือ ข้อมูลมีค่าที่ใกล้เคียงกับค่าของข้อมูลที่พิจารณามากที่สุด ในที่นี้ค่าความใกล้เคียงจะหมายถึง ระยะทางที่มีค่าน้อยที่สุดระหว่างชุดข้อมูลกับข้อมูลที่พิจารณา ข้อมูลดังกล่าวนี้เรียกว่า Nearest Neighbor โดยระยะห่างของข้อมูลนั้นสามารถกำหนดได้ว่าต้องการมากน้อยเพียงใด ด้วยเหตุนี้จึงมีการจำแนกที่เรียกว่า K-NN ซึ่ง K แทนด้วยค่าระยะทางระหว่างข้อมูลที่ต้องการ (Moosavian, et al., 2013)

โดยงานวิจัยนี้ได้นำอัลกอริทึม K-Nearest Neighbor ที่ใช้ระยะทางแฮมมิ่ง (Hamming Distance) มาคำนวณหาความคล้ายคลึงกับชุดข้อมูลท่องเที่ยว โดยมีหลักการคือหาผลรวมของตำแหน่งตัวอักษรที่ไม่เหมือนกัน ทั้งนี้เนื่องจากเป็นวิธีการที่คำนวณได้ง่ายและรวดเร็ว จึงเหมาะสมกับการนำมาใช้ในการแนะนำข้อมูลท่องเที่ยวนี้ เมื่อผู้ใช้เลือกจะให้เป็นตัวเลข 1 หากผู้ใช้ไม่เลือกจะให้เป็นตัวเลข 0 จากนั้นเปรียบเทียบชุดข้อมูลท่องเที่ยวและชุดข้อมูลความต้องการของผู้ใช้ ดังสมการ 8 แล้วแนะนำข้อมูลท่องเที่ยวตามชุดข้อมูลที่คำนวณค่าระยะทางน้อยที่สุดในหมวดสถานที่ท่องเที่ยว ที่พัก และร้านอาหาร ซึ่งผู้ใช้จะต้องป้อนข้อมูลท่องเที่ยวที่สนใจ คือ ประเภทของการท่องเที่ยว ที่พัก ร้านอาหาร สิ่งอำนวยความสะดวก และกิจกรรมท่องเที่ยวที่ผู้ใช้ต้องการ

$$D_H = \sum_{i=1}^{n} \left(x_i \# y_i \right) \quad (8)$$

เมื่อ D_H คือ ผลรวมของชุดข้อมูลท่องเที่ยว x ที่ไม่เท่ากับกับชุดข้อมูลท่องเที่ยวที่ผู้ใช้เลือก y

x คือ ชุดข้อมูลท่องเที่ยว

y คือ ชุดข้อมูลท่องเที่ยวที่ผู้ใช้เลือก

n คือ จำนวนข้อมูลที่ใช้สำหรับพิจารณาทั้งหมด

จากสมการที่ 8 สามารถแสดงตัวอย่างการคำนวณหาระยะทางแฮมมิ่ง โดยกำหนดให้เลข 1 ของชุดข้อมูลท่องเที่ยว หมายถึง ข้อมูลท่องเที่ยวมีข้อมูลเหล่านั้นอยู่และหากเป็น 1 ในชุดข้อมูลผู้ใช้จะหมายถึงผู้ใช้ต้องการข้อมูลท่องเที่ยวประเภทนั้นๆ ซึ่งสามารถคำนวณระยะทางแฮมมิ่งได้ดังต่อไปนี้

ตาราง 6 แสดงตัวอย่างชุดข้อมูลท่องเที่ยวกับชุดข้อมูลที่ผู้ใช้เลือก

	ข้อมูลที่ 1	ข้อมูลที่ 2	ข้อมูลที่ 3	ข้อมูลที่ 4	ข้อมูลที่ 5	ข้อมูลที่ 6
ชุดข้อมูลท่องเที่ยว	1	0	1	0	1	0
ชุดข้อมูลที่ผู้ใช้เลือก	1	1	0	0	0	1
$x_i \# y_i$	0	1	1	0	1	1

จากตาราง 6 ให้พิจารณาชุดข้อมูลท่องเที่ยวกับชุดข้อมูลที่ผู้ใช้เลือก เมื่อชุดข้อมูลทั้งสองมีความเหมือนกันจะกำหนดให้เป็นเลข 0 และหากชุดข้อมูลทั้งสองไม่เหมือนกันจะกำหนดให้เป็นเลข 1 จากนั้นคำนวณหาค่า K-NN ที่ใช้ระยะทางแฮมมิ่ง โดยได้มาจากผลรวมของชุดข้อมูลท่องเที่ยวที่ไม่ตรงกับชุดข้อมูลที่ผู้ใช้เลือก ซึ่งในตัวอย่างนี้จะได้ค่า K-NN ที่ใช้ระยะทางแฮมมิ่งเท่ากับ 4

งานวิจัยที่เกี่ยวข้อง

การใช้อินเทอร์เน็ตเพื่อการสืบค้นข้อมูลได้รับความนิยมอย่างแพร่หลาย เนื่องจากถือได้ว่าเป็นคลังข้อมูลที่ใหญ่ โดยเฉพาะข้อมูลทางด้านการท่องเที่ยวที่มีความสำคัญต่อนักท่องเที่ยวเป็นอย่างมาก โดยมีงานวิจัยที่ได้นำเทคโนโลยีและแนวคิดต่างๆ มาช่วยนำเสนอข้อมูลด้านการท่องเที่ยวให้มีประสิทธิภาพซึ่งในปัจจุบันมีเว็บแนะนำการท่องเที่ยวหรือเว็บที่ให้ข้อมูลการท่องเที่ยวอย่างมากมาย มีทั้งเว็บไซต์ที่แสดงผลแบบข้อความ รูปภาพ ภาพเคลื่อนไหว หรือ ทั้งหมดผสมกัน รวมถึงมีงานวิจัยที่เกี่ยวกับการสืบค้น แนะนำและวางแผนการท่องเที่ยวต่างๆ มากมาย ดังแสดงผลการเปรียบเทียบแต่ละงานในตาราง 7

ตาราง 7 แสดงผลการเปรียบเทียบแต่ละงาน

งานวิจัย	รายละเอียด	ใช้วิธีการ/เทคนิค	แก้ไข/ปรับปรุง/เพิ่มเติม
Dema (2008) eTourPlan	ค้นหาและวางแผนท่องเที่ยวประเทศภูฐาน	ออนโทโลจีเก็บข้อมูลท่องเที่ยวและใช้ OO jDREW สืบค้นข้อมูลท่องเที่ยว	เนื่องจากงานวิจัยนี้ยังไม่ได้สร้างเป็นเว็บไซต์ การใช้งานผ่าน OO jDREW ยุ่งยาก ซึ่งผู้วิจัยได้นำแนวคิดในการออกแบบออนโทโลจีและกฎมาพัฒนาในรูปแบบของเว็บไซต์
Siricharoen (2008) e-Tourism	ในระบบ e-Tourism จะต้องประกอบไปด้วยคลาสต่างๆ คือWhere (เที่ยว/พัก ที่ไหน) When (ไปเมื่อไหร่) What (ทำกิจกรรมอะไร) Transportation (เดินทางไปอย่างไร)	ออกแบบออนโทโลจี	เพิ่มคลาสจังหวัด อำเภอ ตำบล ร้านอาหาร ร้านขายของฝาก ร้านหนึ่งตำบลหนึ่งผลิตภัณฑ์ เทศกาล เพื่อให้ครอบคลุมกับความต้องการของนักท่องเที่ยว

ตาราง 7 (ต่อ)

งานวิจัย	รายละเอียด	ใช้วิธีการ/เทคนิค	แก้ไข/ปรับปรุง/เพิ่มเติม
Castillo, et al. (2008)	วางแผนการท่องเที่ยวแบบไดนามิกและใช้ข้อมูลพฤติกรรมการท่องเที่ยวในอดีตของผู้ใช้และจากความคล้ายคลึงกับผู้ใช้คนอื่น พร้อมแนะนำเส้นทางการเดินทางในรูปแบบไฟล์ JPG หรือ PDF	Case-based Reasoning/ออนโทโลจีเก็บข้อมูลท่องเที่ยว/โทรศัพท์มือถือ/A* แนะนำเส้นทาง	นำแนวคิดมาประยุกต์ใช้ในการออกแบบคำถามแบบตัวเลือก
Choi, et al. (2009)	แนะนำการเดินทางท่องเที่ยวเกาะเชจู ประเทศเกาหลี ประกอบไปด้วย สถานที่ท่องเที่ยว ที่พัก ร้านอาหาร รถเช่าและการเดินทาง รับความคิดเห็นจากผู้ใช้และวิเคราะห์จัดเรทติ้ง	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/พัฒนาเว็บเชิงความหมาย/ Web Crawler ในการดึงข้อมูลท่องเที่ยวที่อยู่ในอินเทอร์เน็ต/ CQEFT เป็น Inference Engine	เพิ่มออนโทโลจีเชิงเวลาที่แนะนำข้อมูลท่องเที่ยวตามฤดูกาลที่เหมาะสมกับสถานที่ท่องเที่ยวและตามปฏิทินจันทรคติที่ในแต่ละปีจะไม่ตรงกัน เพื่อให้ไม่พลาดเทศกาลสำคัญๆ ที่มีในจังหวัดนั้นๆ
Schiaffino and Amandi (2009)	แนะนำการท่องเที่ยวและจัดแพ็คเก็จในวันหยุด	ระบบผู้เชี่ยวชาญ/ Collaborative fitering + Content-based user profiles + Demographic information/ ออนโทโลจีเก็บข้อมูลท่องเที่ยว	มีเฉพาะข้อมูลที่พัก รีสอร์ท และสายการบินเท่านั้น

ตาราง 7 (ต่อ)

งานวิจัย	รายละเอียด	ใช้วิธีการ/เทคนิค	แก้ไข/ปรับปรุง/เพิ่มเติม
Park, et al. (2009)	พัฒนาศูนย์บริการข้อมูลท่องเที่ยวบนโทรศัพท์มือถือ มีรายงานสภาพการจราจร มีแถบ RFID ติดบริเวณสถานที่ท่องเที่ยว สามารถใช้มือถืออ่านข้อมูลท่องเที่ยวได้	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/โทรศัพท์มือถือหรือ PDA	พัฒนาระบบการสืบค้นข้อมูลท่องเที่ยว เนื่องจากงานนี้ไม่มีระบบการค้นหาสถานที่ท่องเที่ยวซึ่งไม่สะดวกในการใช้งาน
Daramola (2010)	แสดงให้เห็นถึงความสำคัญในการใช้ออนโทโลจีกับปัญหาการท่องเที่ยว โดยจัดเก็บข้อมูลท่องเที่ยวต่างๆ คือ สถานที่ท่องเที่ยว ที่พัก เทศกาล ร้านอาหาร และการเดินทาง นำเสนอแนวคิดในการสืบค้นเชิงความหมาย	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/เว็บเชิงความหมาย	เพิ่มการสืบค้นข้อมูลท่องเที่ยวจากอินเทอร์เน็ตและวิเคราะห์เว็บไซต์เพื่อจัดหมวดหมู่เว็บไซต์ท่องเที่ยวด้วยอัลกอริทึมนาอีฟเบย์ที่ปรับปรุงโดยใช้ค่าขอบเขต
Cao, et al. (2011) และ Cao and Nguyen (2012)	พัฒนาระบบ STAAR (Semantic Tourist information Access and Recommending) สำหรับค้นหาและแนะนำข้อมูลท่องเที่ยวจากการป้อนคำถามที่เป็นภาษาอังกฤษ	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/ใช้งานผ่านโทรศัพท์มือถือ/การประมวลผลภาษาธรรมชาติ	เปลี่ยนกระบวนการตัดคำและกระบวนการเทียบจากพจนานุกรมหรือคลังคำศัพท์มาใช้อัลกอริทึม String Matching ในการเข้าคู่คำเพื่อแนะนำข้อมูลท่องเที่ยวในการป้อนคำถามรูปแบบภาษาธรรมชาติ

ตาราง 7 (ต่อ)

งานวิจัย	รายละเอียด	ใช้วิธีการ/เทคนิค	แก้ไข/ปรับปรุง/เพิ่มเติม
Kongthon, et al. (2011)	ป้อนคำถามที่เป็นภาษาไทยเพื่อขอข้อมูลท่องเที่ยว	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/Web Crawler ในการดึงข้อมูลท่องเที่ยวที่อยู่ในอินเทอร์เน็ต/ SPARQL/ การประมวลผลภาษาธรรมชาติ	เปลี่ยนกระบวนการตัดคำและกระบวนการเทียบจากพจนานุกรมหรือคลังคำศัพท์มาใช้อัลกอริทึม String Matching ในการเข้าคู่คำเพื่อแนะนำข้อมูลท่องเที่ยวในการป้อนคำถามรูปแบบภาษาธรรมชาติ
Ruiz-Martinez, et al. (2011)	สกัดความรู้จากเอกสารเว็บและคัดแยกข้อมูลเพื่อจัดเก็บลงในออนโทโลจีตามหมวดหมู่ที่ได้ออกแบบไว้	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/สกัดความรู้ด้วยเทคนิค NERC (Name Entity Recognition and Classification)/ เว็บเชิงความหมาย	เปลี่ยนขั้นตอนการตัดคำเป็นใช้คำที่อยู่ในออนโทโลจีมาเข้าคู่และใช้อัลกอริทึมนาอีฟเบย์ที่ปรับปรุงโดยใช้ค่าขอบเขตในการจัดหมวดหมู่เว็บไซต์ท่องเที่ยว
Garcia-Crespo, et al. (2011)	แนะนำโรงแรม โดยผู้ใช้สามารถป้อนรายละเอียดความต้องการ ได้แก่ บริการที่ตั้ง ฤดูกาลและกิจกรรมต่างๆ และนำค่าเรทติ้งจากความคิดเห็นของผู้ใช้ที่เคยไปพักมาช่วยวิเคราะห์	ระบบผู้เชี่ยวชาญ/ เว็บเชิงความหมาย/ ออนโทโลจี/ ตรรกศาสตร์คลุมเครือ	เพิ่มส่วนการแนะนำสถานที่ท่องเที่ยว ร้านอาหาร ร้านขายของฝาก ที่ใกล้เคียงกับที่พัก และนำแนวคิดในการเลือกที่พักที่มีรายละเอียดย่อยๆ มาใช้ในการแนะนำข้อมูลท่องเที่ยวในรูปแบบของคำถามแบบตัวเลือก

ตาราง 7 (ต่อ)

งานวิจัย	รายละเอียด	ใช้วิธีการ/เทคนิค	แก้ไข/ปรับปรุง/เพิ่มเติม
Moreno, et al. (2011)	แนะนำและวางแผนท่องเที่ยว โดยใช้พฤติกรรมของผู้ใช้มาพิจารณาด้วยเทคนิค Collaborative พร้อมแสดงแผนที่กูเกิ้ลและระบุตำแหน่งพิกัดทางภูมิศาสตร์เพื่อให้ผู้ใช้ไม่หลงทางแต่แนะนำเฉพาะเมืองตาราโกนา ประเทศสเปน ด้วยภาษาสเปนเท่านั้น	Collaborative/เว็บเชิงความหมาย	เพิ่มเติมในส่วนของการดึงพิกัดทางภูมิศาสตร์ของตำแหน่งสถานที่ที่ผู้ใช้สนใจเพื่อแนะนำสถานที่ท่องเที่ยว ร้านอาหาร ร้านขายของฝากใกล้เคียง
Rajkumar, et al. (2012)	แสดงถึงประโยชน์ที่มีการสืบค้นด้วยเว็บเชิงความหมายในธุรกิจท่องเที่ยวที่พบว่าใช้คำค้นหาที่ไม่เหมาะสมทำให้ไม่พบข้อมูลท่องเที่ยวในอินเทอร์เน็ต จุดเด่นคือสามารถเลือกสถานที่ท่องเที่ยวที่จะไปตามลำดับ มีการใช้ข้อมูลอายุ อาชีพ ในการแนะนำกิจกรรมที่เหมาะสมและมีการแสดงสภาพภูมิอากาศของสถานที่ท่องเที่ยว	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/SPARQL/เว็บเชิงความหมาย	เพิ่มอัลกอริทึม Name Variation Matching ในการสร้างคำที่สามารถเขียนได้หลากหลายรูปแบบเพื่อให้เว็บไซต์ที่มีคำเขียนหรือสะกดไม่ถูกต้องตามหลักของภาษาไทยและภาษาอังกฤษถูกค้นพบ แต่เนื่องจากคำที่ถูกสร้างจากอัลกอริทึม Name Variation Matching มีจำนวนมาก ผู้วิจัยจึงใช้อัลกอริทึม ISG มาช่วยกรองคำค้นหาเพื่อไม่ให้มีปริมาณข้อมูลมากเกินความจำเป็น และนำแนวคิดในการใช้อายุมาช่วยกำหนดกฎในออนโทโลจีเพื่อแนะนำสถานที่ท่องเที่ยวให้เหมาะสมกับวัย และใช้อัลกอริทึม K-NN ที่ปรับปรุงในการแนะนำข้อมูลท่องเที่ยวตามความต้องการของผู้ใช้จากคำถามแบบตัวเลือก

ตาราง 7 (ต่อ)

งานวิจัย	รายละเอียด	ใช้วิธีการ/เทคนิค	แก้ไข/ปรับปรุง/เพิ่มเติม
Palaniammal, et al. (2012)	แนะนำการท่องเที่ยวที่ให้รายละเอียดเกี่ยวกับที่พัก ที่ช้อปปิ้ง สถานที่ท่องเที่ยว ร้านอาหารและการเดินทาง โดยผู้ใช้สามารถเลือกสถานที่ท่องเที่ยวที่สนใจจะไปเที่ยวตามลำดับ อายุอาชีพ เพื่อสร้างกิจกรรมที่เหมาะสม	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/ SPARQL	นำแนวคิดมาใช้ในการออกแบบคำถามในรูปแบบตัวเลือกเพื่อแนะนำข้อมูลท่องเที่ยวให้เหมาะสมและตรงความต้องการของผู้ใช้มากที่สุด
Chaves, et al. (2012)	แนะนำที่พัก โดยมีรายละเอียดสิ่งอำนวยความสะดวกต่างๆ ได้แก่ ภายใน ภายนอก ห้องพัก ห้องน้ำ คนขับ ผู้พิการ รถรับส่งสนามบิน ที่นั่งเด็กเล็ก เช่าจักรยาน เช่ารถยนต์ เข้าพัก ชำระเงิน ผู้ช่วยส่วนตัว แลกเงิน ซักรีด หนังสือพิมพ์ สำหรับคู่แต่งงานใหม่ ฝากกุญแจ นำเที่ยว	ออนโทโลจีเก็บข้อมูลที่พัก	นำแนวคิดในเรื่องสิ่งอำนวยความสะดวกของที่พักมาประยุกต์ใช้ในการออกแบบคำถามในรูปแบบตัวเลือกและปรับปรุง เพิ่มเติมรายละเอียดให้เหมาะสมกับการท่องเที่ยวประเทศไทย เช่น มีสระว่ายน้ำ มีห้องอบซาวน่า สปา มีบริการนวดแผนไทย เป็นต้น
Ananthapadmanaban, et al. (2012)	การสืบค้นข้อมูลท่องเที่ยว ได้แก่ สถานที่ท่องเที่ยว ที่พัก โรงพยาบาลหรือศูนย์สุขภาพ กิจกรรมและการเดินทางในรัฐทมิฬนาฑู ประเทศอินเดียผ่านเว็บ	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/ Web Service/เว็บเชิงความหมาย/ SPARQL/Web Crawler เก็บข้อมูลท่องเที่ยวจากอินเทอร์เน็ต	สามารถใช้เป็นแนวทางในงานวิจัยได้ ซึ่งผู้วิจัยได้เพิ่มเติมในส่วนของเทศกาลและกิจกรรมที่พบในสถานที่ท่องเที่ยว ที่พัก ร้านอาหาร เพื่อให้ครอบคลุมกับความต้องการของนักท่องเที่ยว

ตาราง 7 (ต่อ)

งานวิจัย	รายละเอียด	ใช้วิธีการ/เทคนิค	แก้ไข/ปรับปรุง/เพิ่มเติม
	เซอร์วิสเชิงความหมาย รวมถึงแนะนำสถานที่ท่องเที่ยวใกล้เคียงโดยการสร้างความสัมพันธ์ของข้อมูลในออนโทโลจี		
Hristoskova, et al. (2012)	แนะนำการท่องเที่ยวอาคาร เช่น พิพิธภัณฑ์ ธนาคาร ศูนย์วิทยาศาสตร์	ออนโทโลจีเก็บข้อมูลท่องเที่ยว/เขียนกฎด้วยภาษา SWRL	นำแนวคิดในการสร้างกฎด้วยภาษา SWRL มาใช้และสร้างกฎให้เหมาะสมกับออนโทโลจีที่ผู้วิจัยได้ออกแบบไว้
จุฑามาศ ศิริรัชนีกร และ ฤทธิกร เล้าอรุณ (2549)	ระบบค้นหาข้อมูลแบบซับซ้อนสำหรับข้อมูลการท่องเที่ยวในประเทศไทย	เว็บเชิงความหมาย/SPARQL	นำเสนอข้อมูลเฉพาะภาษาอังกฤษเท่านั้น ซึ่งผู้วิจัย ได้นำแนวคิดมาพัฒนาให้สามารถสืบค้นได้ทั้งภาษาไทยและภาษาอังกฤษ
สุรางค์รัตน์ เชาว์โคกสูง และ จักรกฤษณ์ เสน่ห์ (2552)	ระบบบริหารจัดการข้อมูลและเชื่อมโยงเส้นทางแหล่งท่องเที่ยวแบบอนิเมชั่นเพื่อนำเสนอในรูปแบบที่น่าสนใจและช่วยดึงดูดนักท่องเที่ยวให้เข้ามาค้นหาข้อมูล	อนิเมชั่น	นำแนวคิดมาเพิ่มเติมให้ครอบคลุมครบทั้ง 77 จังหวัดในประเทศไทย ซึ่งงานนี้นำเสนอข้อมูลท่องเที่ยวจังหวัดพิษณุโลก สุโขทัย เพชรบูรณ์ และอุตรดิตถ์เท่านั้น

อย่างไรก็ตามยังพบว่ามีการพัฒนาเว็บไซต์ที่นำเสนอและแนะนำข้อมูลท่องเที่ยวในประเทศไทย ดังเช่น การท่องเที่ยวแห่งประเทศไทย (http://thai.tourismthailand.org) ได้พัฒนาเว็บไซต์ที่นำเสนอข้อมูลท่องเที่ยวประเทศไทยในรูปแบบภาษาไทย ภาษาอังกฤษและภาษาอื่นๆ โดยให้บริการข้อมูลท่องเที่ยวอันได้แก่ ประเทศไทย การเดินทาง ข้อมูลจังหวัด สถานที่ท่องเที่ยวและกิจกรรม ที่พัก ร้านอาหาร ช้อปปิ้ง โปรโมชั่น ประสบการณ์นักเดินทางและข่าวอัพเดท แต่ยังมีข้อด้อยต่างๆ ต่อไปนี้

1. ระบบทั่วๆ ไปยังไม่สามารถรองรับการดึงรายละเอียดข้อมูลเว็บไซต์ที่สะกดผิดภาษาไทยและภาษาอังกฤษ

2. ระบบยังไม่รองรับการค้นหาในกรณีสถานที่นั้นๆ มีชื่อเรียกหลากหลาย

3. ไม่มีระบบแนะนำข้อมูลท่องเที่ยวที่ใกล้เคียงกับสถานที่ท่องเที่ยวนั้นๆ

ไทยแลนด์แมพไกด์ (http://www.thailand-map-guide.com) ได้พัฒนาระบบแนะนำเส้นทางการเดินทางไปยังสถานที่ต่างๆ โดยใช้แผนที่จากกูเกิ้ล แสดงระยะทางในการเดินทาง ความเร็วที่สามารถใช้ได้และคำอธิบายเส้นทาง พร้อมทั้งแสดงการเดินทางเป็นภาพเคลื่อนไหวเป็นรูปรถยนต์ไปตามเส้นทางที่ระบบแนะนำ

ข้อดีของระบบคือสามารถนำแผนที่ไปวางในโปรแกรม Microsoft Word แต่ระบบนี้ยังมีข้อเสียที่ระบบไม่รองรับการระบุสถานที่จุดเริ่มต้นและปลายทางทำให้ไม่สะดวกในการใช้งานมากนัก อีกทั้งยังเสียเวลาในการเลือกสถานที่ที่ต้องการและไม่มีส่วนสำหรับการป้อนคำค้นหาสถานที่ที่ต้องการเดินทางไปท่องเที่ยว

ไทยโรเวอร์เนร์ธ (http://www.thailowernorth.com) เป็นเว็บไซต์ที่นำเสนอข้อมูลท่องเที่ยวกลุ่มจังหวัดภาคเหนือตอนล่างประกอบด้วย จังหวัดนครสวรรค์ จังหวัดอุทัยธานี จังหวัดกำแพงเพชรและจังหวัดพิจิตร โดยมีข้อมูลเกี่ยวกับจังหวัด สถานที่ท่องเที่ยว ที่พัก ที่กิน ของฝาก ปฏิทินท่องเที่ยว การเดินทาง เทศบาลประจำจังหวัดและหน่วยงานที่สำคัญ ซึ่งข้อมูลท่องเที่ยวถูกนำเสนอในรูปแบบข้อความ รูปภาพและแสดงตำแหน่งของสถานที่นั้นๆ ด้วยแผนที่จากกูเกิ้ล

ข้อดีของระบบคือสามารถใช้เป็นแหล่งอ้างอิงข้อมูลท่องเที่ยวได้ แต่หากต้องการค้นหาสถานที่ท่องเที่ยว ที่พัก ที่กินจะต้องคลิกเข้าไปในหัวข้อแต่ละจังหวัดไม่สามารถค้นหาในหน้าแรกได้อีกทั้งระบบค้นหายังไม่รองรับการพิมพ์ชื่อหรือสะกดพิมพ์ในภาษาไทยและภาษาอังกฤษ ไม่รองรับการค้นหาในกรณีสถานที่นั้นๆ มีชื่อเรียกหลากหลายและไม่สามารถเชื่อมโยงข้อมูลท่องเที่ยวบริเวณใกล้เคียงกันได้

จากข้อวิจารณ์และข้อจำกัดในงานวรรณกรรมเหล่านี้เป็นข้อเสนอแนะที่สำคัญอย่างยิ่งในการนำมาปรับปรุงอัลกอริทึมเนมแมทชิ่ง นาอีฟเบย์และออนโทโลจีเชิงเวลาสำหรับระบบแนะนำข้อมูลท่องเที่ยวในประเทศไทย โดยผู้วิจัยประยุกต์ใช้ทฤษฎี หลักการและเทคนิคต่างๆ ในงานวรรณกรรมดังต่อไปนี้

1. สำหรับการออกแบบและพัฒนาออนโทโลจีท่องเที่ยว ได้นำแนวคิดการท่องเที่ยวภูฐาน (Dema, 2008) มาประยุกต์ใช้ในการออกแบบและพัฒนาออนโทโลจีท่องเที่ยวในประเทศไทย (นฤพนธ์ พนาวงศ์ และจักรกฤษณ์ เสน่ห์, 2553ก, 2553ข) รวมถึงกำหนดความสัมพันธ์และ

เชื่อมโยงข้อมูลท่องเที่ยวเพื่อให้ง่ายต่อการสืบค้น ซึ่งออนโทโลจีท่องเที่ยวนี้จะถูกจัดเก็บเป็นไฟล์รูปแบบ OWL

2. สำหรับการจัดหมวดหมู่เว็บไซต์ท่องเที่ยว ในส่วนนี้ผู้วิจัยประยุกต์ใช้อัลกอริทึมเนมแมท-ซิ่งที่ใช้อัลกอริทึม ISG (Snae, 2006) ในการกรองข้อมูลชื่อที่สามารถเขียนได้หลายรูปแบบหรือมีความคล้ายคลึงกัน ทั้งคำพ้องรูป พ้องเสียงหรือสะกดผิดในภาษาไทยหรือภาษาอังกฤษ เพื่อให้เว็บไซต์เหล่านั้นถูกค้นพบ จากนั้นนำชื่อที่สามารถเขียนได้หลากหลายไปสืบค้นข้อมูลท่องเที่ยวในเว็บไซต์กูเกิ้ลด้วยอัลกอริทึมของ Web Crawler (นฤพนธ์ พนาวงศ์ และจักรกฤษณ์ เสน่ห์ นมะหุต, 2556ก) แล้วนำข้อมูลท่องเที่ยวมาจัดหมวดหมู่เว็บไซต์ 6 หมวดหมู่ คือ สถานที่ท่องเที่ยว ที่พัก ร้านอาหาร ร้านขายของฝาก ร้านสินค้าหนึ่งตำบลหนึ่งผลิตภัณฑ์และเทศกาลด้วยอัลกอริทึม LSI ซึ่งผู้วิจัยนำแนวคิดมาจากงานวิจัยของ Thorleuchter and Poel (2013) อัลกอริทึมนาอีฟเบย์ (Patil and Pawar, 2012) และทำการปรับปรุงอัลกอริทึมนาอีฟเบย์โดยใช้ค่าขอบเขต

3. สำหรับการพัฒนาระบบแนะนำข้อมูลท่องเที่ยวในประเทศไทย แบ่งเป็น 2 ส่วนคือ

3.1 จากการสืบค้นข้อมูลท่องเที่ยวโดยใช้คำค้นหาหรือใช้ประโยค ซึ่งผู้วิจัยใช้ภาษา SPARQL สำหรับสืบค้นข้อมูลออนโทโลจีท่องเที่ยวในประเทศไทย (นฤพนธ์ พนาวงศ์ และจักรกฤษณ์ เสน่ห์, 2554) เกี่ยวกับจังหวัดที่ต้องการไปท่องเที่ยว หรือ ใช้คำค้นหาที่อาจหมายถึงจังหวัดนั้นๆ เช่น ชื่อเรียกเดิมของจังหวัด คำขวัญประจำจังหวัด สัญลักษณ์ประจำจังหวัด ต้นไม้ประจำจังหวัด ดอกไม้ประจำจังหวัด เป็นต้น โดยใช้ Jena API ทำหน้าที่เป็น OWL Parser เพื่ออ่านโครงสร้างของข้อมูลท่องเที่ยวในออนโทโลจีท่องเที่ยวตามที่ได้ออกแบบและพัฒนาไว้แล้วในข้อที่ 1

3.2 จากการเลือกข้อมูลท่องเที่ยวตามความต้องการของผู้ใช้ ซึ่งผู้วิจัยใช้ภาษา SWRL ในการสร้างกฎเกี่ยวกับการแนะนำข้อมูลท่องเที่ยวในประเทศไทย ประกอบด้วย 5 ตัวแปร คือ นักท่องเที่ยว อายุ เพศ สถานที่ท่องเที่ยวและจังหวัดที่ต้องการไปเที่ยว โดยใช้ Protégé API 3.5 ในการเชื่อมต่อกับ Jess Inference Engine เพื่ออนุมานข้อมูลท่องเที่ยวประจำจังหวัดที่ต้องการไปเที่ยวให้เหมาะสมกับเพศและอายุของผู้ใช้ รวมถึงการใช้อัลกอริทึมในการแนะนำข้อมูลให้ตรงตามความต้องการและตามฤดูกาลหรือช่วงเวลาที่เหมาะสมกับการท่องเที่ยว ดังรายละเอียดต่อไปนี้

3.2.1 สำหรับอัลกอริทึม K-NN (Moosavian, et al., 2013) ถูกนำไปใช้เพื่อแนะนำข้อมูลท่องเที่ยวในประเทศไทยให้ตรงตามความต้องการของผู้ใช้ โดยผู้ใช้จะต้องเลือกประเภทสถานที่ท่องเที่ยว ประเภทที่พัก ประเภทอาหาร สิ่งอำนวยความสะดวกและกิจกรรมที่สนใจท่องเที่ยวของจังหวัดที่ต้องการไปเที่ยว

3.2.2 ในการพัฒนาอัลกอริทึมสำหรับคำนวณปฏิทินจันทรคติโดยใช้ออนโทโลจีเชิงเวลา (นฤพนธ์ พนาวงศ์ และจักรกฤษณ์ เสน่ห์ นมะหุต, 2556ข) เพื่อแนะนำข้อมูลท่องเที่ยวตามช่วงเวลาให้เหมาะสมกับฤดูกาลท่องเที่ยวที่มีในประเทศไทยและช่วยแนะนำข้อมูลท่องเที่ยวตามวันสำคัญต่างๆ ที่ในแต่ละปีจะไม่ตรงกัน