

ระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยด้วยหลักการออนโทโลยีและเนมแมต칭

Search System for Attractions in Thailand with Ontology and Name Matching

นฤพนธ์ พนาวงศ์ และ จักรกฤษณ์ เสน่ห์*

ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์

มหาวิทยาลัยนเรศวร จังหวัดพิษณุโลก 65000

*Corresponding author. E-mail: chakkrits@nu.ac.th

บทคัดย่อ – การณรงค์การท่องเที่ยวในประเทศไทยมีการกระตุ้นและมีการพัฒนาอย่างต่อเนื่องเพื่อให้เศรษฐกิจไทยฟื้นตัวและเพิ่มมูลค่าทางเศรษฐกิจ ปัจจุบันมีระบบต่าง ๆ ที่จะช่วยให้ข้อมูลต่าง ๆ กับนักท่องเที่ยวไม่ว่าจะอยู่ในรูปแบบสื่อมัลติมีเดียหรือเว็บไซต์ต่าง ๆ แต่การที่นักท่องเที่ยวจะสืบค้นในการไปเที่ยวสถานที่แห่งใดแห่งหนึ่งนั้น นักท่องเที่ยวจำเป็นต้องรู้สถานที่ท่องเที่ยว ที่พัก ร้านอาหาร ร้านขายของฝาก ร้านสินค้าหนึ่งตำบลหนึ่งผลิตภัณฑ์ของบริเวณหรือจังหวัดนั้น ๆ ซึ่งระบบการค้นหาสถานที่ท่องเที่ยวโดยผ่านเสิร์ชเอนจินต่าง ๆ ทั่วไปนั้นจะให้ผลลัพธ์เฉพาะเรื่องที่ต้องการค้นหาอย่างเดียวทำให้ต้องเสียเวลาในการเข้าถึงข้อมูลจากหลายเว็บไซต์เพื่อให้ได้ผลลัพธ์ตรงความต้องการทั้งหมด ดังนั้นผู้วิจัยจึงได้นำหลักการออนโทโลยีมาประยุกต์ใช้ในการพัฒนาระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยเพื่อให้ข้อมูลออนไลน์ที่แตกต่างกันในเนื้อหาเดียวกันโดยสามารถเชื่อมโยงข้อมูลการท่องเที่ยวต่าง ๆ อย่างมีความสัมพันธ์ต่อกัน อีกทั้งผู้วิจัยได้ใช้ทฤษฎีเนมแมต칭มาพัฒนาอัลกอริทึมในการค้นหาคำที่มีความหลายหลายและเขียนผิดหรือไม่ถูกต้องตามหลักภาษาไทยและภาษาอังกฤษที่สามารถค้นเจอในเว็บไซต์ต่าง ๆ ได้อีกด้วย และยังช่วยประชาสัมพันธ์แหล่งท่องเที่ยวที่มีการเขียนผิดในอินเทอร์เน็ตสามารถถูกค้นพบได้และได้ผลลัพธ์ที่ตรงกับความต้องการในการค้นหาคำตอบในครั้งเดียว

คำสำคัญ – การท่องเที่ยว, ออนโทโลยี, การค้นหา, เนมแมต칭, การเข้าถึงข้อมูล

ABSTRACT– The campaign for economic recovery in Thailand is encouraged in several areas, one of them the tourism industry. There are currently systems that will show information of various kinds to tourists, whether in the textual form or as multimedia sites, which the visitors are searching for places to visit. Tourists need to know various details, e.g. attractions, accommodation, restaurants, souvenir shops and OTOP shops (One Subdistrict One Product). When tourists are using a search engine they will most likely end up by visiting different Web sites for all details, which makes the process of gathering all the necessary data laborious. In this paper, we present a way to improve this situation by introducing an ontology for develop of search systems for attraction in Thailand. Moreover, we use a name matching algorithm that can deal with name variations both in Latin and Thai scripts. This algorithm ensures that misspelled or otherwise incorrectly written names can still be used and found by the search engine. With our system erroneous spellings also can be found on the Internet easily.

KEY WORDS – tourism, ontology, search, name matching, information retrieval

1. บทนำ

ปัจจุบันการท่องเที่ยวถือเป็นอุตสาหกรรมบริการที่ทำรายได้ให้แก่ประเทศต่าง ๆ เป็นจำนวนมาก รวมทั้งประเทศไทย โดยมีบทบาทอย่างมากต่อการพัฒนาทางด้านเศรษฐกิจภายในประเทศ เนื่องจากเป็นอุตสาหกรรมที่มีการเติบโตอย่างต่อเนื่อง รวมทั้งได้มีการพัฒนาแหล่งท่องเที่ยว วิธีการจัดการการท่องเที่ยวให้ขยายตัวมากขึ้น ซึ่งจะเป็นปัจจัยที่สามารถสร้างเศรษฐกิจหมุนเวียนและเป็นการกระจายรายได้ไปสู่ท้องถิ่นภายในประเทศเป็นอย่างดี จากการส่งเสริมและพัฒนาการท่องเที่ยวอย่างต่อเนื่องของการท่องเที่ยวแห่งประเทศไทยและสำนักพัฒนาการท่องเที่ยว การท่องเที่ยวในประเทศไทยแบ่งรูปแบบการท่องเที่ยวตามทรัพยากรท่องเที่ยวออกเป็น 2 ลักษณะคือ ทรัพยากรการท่องเที่ยวที่เกิดขึ้นเองตามธรรมชาติ และทรัพยากรการท่องเที่ยวที่มนุษย์สร้างขึ้น [1] ซึ่งจากสถานะเศรษฐกิจในประเทศกำลังฟื้นตัวจากภาวะผลกระทบที่เศรษฐกิจโลกกำลังหดตัว, การเมืองภายในประเทศ, คู่แข่งจากประเทศเพื่อนบ้าน รัฐบาลมีนโยบายสนับสนุนนำโดยการท่องเที่ยวแห่งประเทศไทยได้มีการส่งเสริมและทำการประชาสัมพันธ์การท่องเที่ยวของประเทศไทยให้กับนักท่องเที่ยวไทยและชาวต่างประเทศอย่างต่อเนื่องผ่านทางเว็บไซต์, แผ่นพับ, โทรทัศน์และสื่อต่าง ๆ อีกมากมาย ซึ่งในปี 2552 นี้การท่องเที่ยวแห่งประเทศไทยได้จัดโครงการ “เที่ยวไทยครีกครั้ง เศรษฐกิจไทยคึกคัก” เพื่อมุ่งหวังให้เศรษฐกิจไทยโตขึ้นอีก 5% [2]

โดย [3] ได้ออกแบบสถาปัตยกรรมของระบบเพื่อจัดการข้อมูลองค์กรและบริษัทในจังหวัดพิษณุโลก ซึ่งได้นำ IT-TELL [4] มาช่วยในระบบการค้นหาจากผู้ใช้ที่ใช้ภาษาต่างกัน รวมถึง [5] ยังใช้ออนโทโลยีในการออกแบบเพื่อแก้ปัญหาชื่อ สถานที่ และเวลาให้อยู่ในรูปแบบมาตรฐานเดียวกัน แล้วใช้ทฤษฎีเนมแมทชิงเป็นอัลกอริทึมในการแก้ปัญหาชื่อที่พิมพ์ผิด ซึ่งสามารถนำไปใช้เป็นต้นแบบเพื่อใช้เก็บข้อมูลแหล่งท่องเที่ยวที่พัก ร้านอาหาร ร้านขายของฝาก ร้านขายสินค้าหนึ่งตำบลหนึ่งผลิตภัณฑ์ในประเทศไทยได้ รวมถึง [6] ได้นำเสนอการออกแบบระบบการค้นหาสถานที่ต่าง ๆ ในแผนที่กูเกิ้ล โดยได้นำ Canonical Name Matching มาช่วยในการแก้ปัญหาการพิมพ์ชื่อที่หลากหลาย อีกทั้ง [7] ยังได้นำเสนอระบบการค้นหาข้อมูลแบบซับซ้อนสำหรับข้อมูลการท่องเที่ยวในประเทศไทยที่ใช้เว็บเชิงความหมายที่นำเสนอแต่ข้อมูลภาษาอังกฤษเท่านั้น

อีกทั้งยังไม่รองรับการค้นหาแบบพิมพ์คำผิดที่อาจเกิดจากความไม่ตั้งใจหรือใช้คำที่ไม่ถูกหลักภาษาอังกฤษ

ดังนั้นผู้วิจัยจึงได้พัฒนาระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยด้วยหลักการออนโทโลยีมาออกแบบข้อมูลท่องเที่ยวต่าง ๆ สำหรับการค้นหาเพื่อให้ได้ผลลัพธ์ตรงและครอบคลุมข้อมูลท่องเที่ยวมากที่สุด อีกทั้งยังใช้ทฤษฎีเนมแมทชิงมาแก้ปัญหาเกี่ยวกับคำที่มีการเขียนผิดหรือไม่ถูกต้องตามหลักภาษาไทยและภาษาอังกฤษที่มีอยู่มากมายในเว็บไซต์ท่องเที่ยวต่าง ๆ เพื่อที่จะทำให้มีความสมบูรณ์และระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยสามารถค้นหาเว็บไซต์ต่าง ๆ เหล่านั้นได้อย่างมีประสิทธิภาพ โดยอธิบายวิธีการและการออกแบบออนโทโลยีและออนโทโลยีที่ใช้ในการท่องเที่ยวปัจจุบัน รวมถึงอธิบายทฤษฎีเนมแมทชิงในส่วนที่ 2 และผลของการออกแบบและพัฒนาระบบที่อธิบายไว้ในส่วนที่ 3

2. วิธีการและการออกแบบ

การออกแบบระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยด้วยหลักการออนโทโลยีและเนมแมทชิงนั้นประกอบแบ่งเป็น 2 ส่วนคือ ออนโทโลยีมาใช้เก็บข้อมูลท่องเที่ยวในประเทศไทยเพื่อให้ข้อมูลแสดงอย่างมีความสัมพันธ์กันอย่างถูกต้องครบถ้วนและใช้ทฤษฎีเนมแมทชิงในการแก้ปัญหาจากการค้นหาชื่อจังหวัดที่อาจจะมีการเขียนหรือสะกดคล้ายคลึงกันหรือที่เกี่ยวข้องสามารถค้นหาเจอได้ครบถ้วนมากที่สุดดังแสดงในหัวข้อที่ 2.1 และ 2.2 อีกทั้งยังได้ออกแบบสถาปัตยกรรมของระบบดังแสดงในหัวข้อที่ 2.3

2.1 ออนโทโลยี

ออนโทโลยีได้กลายมาเป็นที่รู้จักและใช้กันอย่างแพร่หลายในปัจจุบัน เช่น สาขาปัญญาประดิษฐ์ สาขาวิชาคอมพิวเตอร์และสาขาวิทยาศาสตร์สารสนเทศ โดยเฉพาะอย่างยิ่งในขอบเขตสาขาทางด้านระบบสารสนเทศ การรวบรวมสารสนเทศทางปัญญา การสืบค้นข้อมูล การแทนความรู้ การออกแบบฐานข้อมูล การจัดการพื้นฐานองค์ความรู้ และระบบการจัดการฐานข้อมูล [8, 9, 10]

[11] ได้นิยามความหมายของออนโทโลยีซึ่งเป็นที่รู้จักกันอย่างกว้างขวางมากที่สุดว่าออนโทโลยีเป็นรายละเอียดที่ชัดเจนแน่นอนของแนวความคิด (“An Ontology is an Explicit Specification of a Conceptualization”) ดังนั้นเมื่อเราต้องการ

แสดงหรือระบุแนวความคิดของวัตถุหรือสิ่งของนั้น ๆ เราควร
จะแสดงรายละเอียดที่ชัดเจนและแน่นอน เช่น การอธิบายออน
โทโลยีของโปรแกรม โดยการนิยามกลุ่มของคำที่แสดงและ
ความสัมพันธ์ระหว่างกลุ่มคำนั้น ๆ การนิยามนี้อาจจะเกี่ยวข้อง
กับชื่อของสิ่งที่มีอยู่ในโปรแกรมซึ่งอาจประกอบไปด้วย
Classes, Relations, Functions หรือ Objects อื่น ๆ

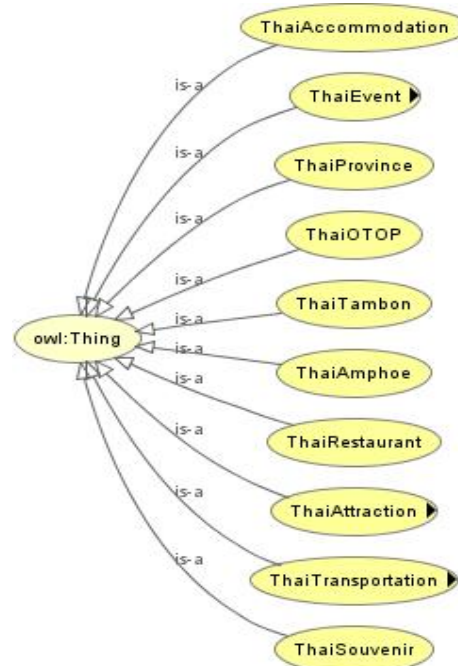
จากนั้น [12] ได้อธิบายคำว่าออนโทโลยีว่าเป็นการแสดง
คำศัพท์และรายละเอียดของขอบเขตหรือวัตถุหรือสิ่งของนั้น ๆ
ไม่เพียงแต่คำศัพท์เท่านั้นแต่แนวคิดที่คำในคำศัพท์มีแนวโน้ม
ที่จะถูกอธิบายและกล่าวถึง ในการออกแบบระบบคอมพิวเตอร์
อาจจะใช้ออนโทโลยีเป็นตัวกำหนดขอบเขตกลไกของ
คอมพิวเตอร์ เช่น ออนโทโลยีอาจจะใช้ระบุแนวคิดของ
ส่วนประกอบของคอมพิวเตอร์ เช่น ซีพียู รอม ดิสค์ไดรฟ์ เมม
โมรี่บอร์ด เป็นต้น และความสัมพันธ์ระหว่างส่วนประกอบของ
คอมพิวเตอร์เหล่านี้ เช่น กลุ่มของกลไกอันหนึ่งอาจจะเป็นชนิด
ย่อยหรือส่วนหนึ่งของอีกสิ่งหนึ่งก็ได้ ในการระบุแนวคิดทั่ว ๆ
ไปนั้น จำเป็นจะต้องใช้การวิเคราะห์ที่อย่างพิถีพิถันของชนิด
ของวัตถุและความสำคัญที่ยังคงมีในขอบเขตนั้น ๆ ซึ่งเรา
เรียกว่า อัปเปอร์ออนโทโลยี (Upper ontology) ซึ่งก็คือออน
โทโลยีที่อธิบายความรู้พื้นฐานทั่ว ๆ ไปที่สามารถจัดการ
อธิบายวัตถุหรือสิ่งของนั้นข้ามไปยังหลาย ๆ สาขาและขอบเขต
ได้

[13] ได้นิยามออนโทโลยีให้เป็นสาขาหนึ่งของวิชา
ปรัชญาที่จัดการเกี่ยวกับชนิดและโครงสร้างของวัตถุ
คุณสมบัติ กระบวนการ และความสัมพันธ์ในทุก ๆ สาขา และ
ขอบเขตของความเป็นจริง เช่น ออนโทโลยีพยายามที่จะระบุว่า
คุณสมบัติอะไรที่ยังคงมีอยู่ในการระบุวัตถุหรือสิ่งของนั้น ๆ
ในทางปัญญาประดิษฐ์และการแทนความรู้และ [14] ก็นิยามว่า
คำว่าออนโทโลยีคือรายละเอียดที่ประกาศอย่างเป็นทางการ
(Formal explicit description) ซึ่งรายละเอียดนั้นถูกแสดงด้วย
คอนเซ็ปต์ต่าง ๆ ในโดเมนที่สามารถกำหนดการอธิบายได้ด้วย
คลาส (Class) สล็อต (Slot) และเงื่อนไข (Restriction) ของ
สล็อตที่เรียกว่าฟาเซต (Facet)

ในปัจจุบันมีการนำออนโทโลยีมาใช้ในการจัดเก็บข้อมูล
การเดินทางและการท่องเที่ยวเพื่อให้เกิดความเชื่อมโยงระหว่าง
ข้อมูลต่าง ๆ เช่น [15] ได้นำเสนอเว็บเชิงความหมายที่ใช้ออน
โทโลยีจัดการการท่องเที่ยววันหยุดในรูปแบบแฟ้มที่มุ่งเน้น
ถึงความสำคัญของธุรกิจขนาดเล็กภายใต้โครงการชื่อ ANOTA

ต่อมา [16] ได้นำเสนอการใช้เว็บเชิงความหมายในธุรกิจ
ท่องเที่ยว ที่ [17] สรุปแนวทางแห่งความสำเร็จและ
องค์ประกอบของธุรกิจท่องเที่ยวบนเครือข่ายอินเทอร์เน็ต (e-
Tourism) ซึ่งประกอบไปด้วยคลาส Where (เที่ยว/พัก ที่ไหน)
When (ไปเมื่อไหร่) What (ทำกิจกรรมอะไร) Transportation
(เดินทางไปอย่างไร) ซ้อนหลังไป 20 ปี ซึ่งสามารถนำมา
ประยุกต์ใช้กับการออกแบบออนโทโลยีของงานวิจัยนี้ได้ อีก
ทั้ง [18] ได้พัฒนาฐานความรู้แหล่งท่องเที่ยวในประเทศฐาน
โดยได้ออกแบบออนโทโลยี และนำ OO jDREW [19] ที่เป็น
ซอฟต์แวร์รหัสเปิด (Open Source) มาใช้ในระบบโดยใช้
เทคโนโลยีจาวาในการพัฒนาเพื่อให้นักท่องเที่ยวสามารถวางแผน
การท่องเที่ยวในประเทศฐานได้ แต่ยังไม่ได้นำเสนอใน
รูปแบบเว็บและไม่ได้ใช้แผนที่ระบบทางภูมิศาสตร์ในการ
นำเสนอแผนที่การเดินทาง เป็นต้น

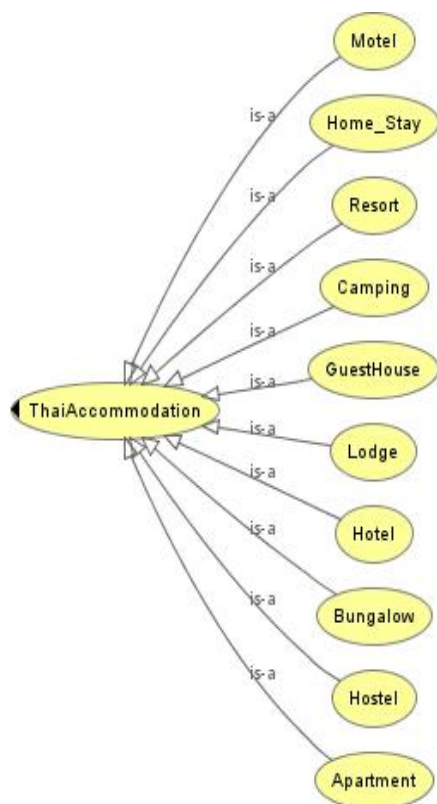
ในงานวิจัยนี้ได้ออกแบบออนโทโลยีท่องเที่ยวที่ใช้
สำหรับเก็บข้อมูลท่องเที่ยวในประเทศไทยดังแสดงในรูปที่ 1



รูปที่ 1. แสดงออนโทโลยีท่องเที่ยว

งานวิจัยนี้ได้ออกแบบออนโทโลยีท่องเที่ยวที่สร้างด้วย
โปรแกรม Protégé 4.1 [20] แบ่งเป็น 10 คลาสคือ คลาสจังหวัด
(ThaiProvince) คลาสอำเภอ (ThaiAmphoe) คลาสตำบล
(ThaiTambon) คลาสเทศกาล (ThaiEvent) คลาสสถานที่

ท่องเที่ยว (ThaiAttraction) คลาสที่พัก (ThaiAccommodation) คลาสการเดินทาง (ThaiTransportation) คลาสร้านอาหาร (ThaiRestaurant) คลาสร้านขายของฝาก (ThaiSouvenir) และ คลาสร้านขายสินค้าหนึ่งผลิตภัณฑ์หนึ่งตำบล (ThaiOTOP) โดยในแต่ละคลาสสามารถแยกรายละเอียดได้ดังแสดงตัวอย่างในรูปที่ 2, 3 และ 4



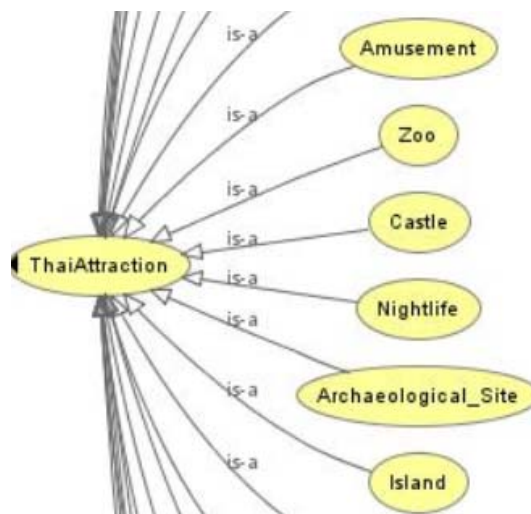
รูปที่ 2. แสดงคลาสที่พัก

จากรูปที่ 2 แสดงการออกแบบรายละเอียดของคลาสที่พัก (ThaiAccommodation) หรือประเภทหรือชนิดที่พักต่าง ๆ ที่มีอยู่ในประเทศไทยที่นักท่องเที่ยวสามารถสืบค้นได้ ได้แก่ โรงแรม (Hotel) รีสอร์ท (Resort) ตั้งแคมป์ (Camping) บ้านเช่าที่พักชั่วคราว (GuestHouse) หอพัก (Hostel) อพาร์ทเมนต์ (Apartment) โฮมสเตย์ (Home_Stay) และบังกะโล (Bungalow) เป็นต้น



รูปที่ 3. แสดงตัวอย่างของคลาสเทศกาล

จากรูปที่ 3 แสดงตัวอย่างของคลาสเทศกาล (ThaiEvent) หรือกิจกรรมต่าง ๆ ที่สามารถเข้าร่วมได้ในประเทศไทย ซึ่งประกอบด้วยสถาปัตยกรรม (Architecture) กีฬา (Sport) ภาพยนตร์ (Theatre) ศิลปะ (Painting) เกษตรกรรม (Agriculture) เทศกาลประจำปี (Annual_Festival) นิทรรศการ (Exhibition) แฟชั่น (Fashion) การออกแบบ (Design) ดูนก (Bird_watching) ถ่ายภาพ (Photography) และปีนเขา (Mountain_Climbing) เป็นต้น



รูปที่ 4. แสดงตัวอย่างของคลาสแหล่งท่องเที่ยว

จากรูปที่ 4 แสดงตัวอย่างของคลาสแหล่งท่องเที่ยวหรือประเภทของแหล่งท่องเที่ยวที่มีอยู่ในประเทศไทยซึ่งประกอบด้วยภูเขา (Mountain) น้ำตก (Waterfall) พืชภัณฑ์

(Museum) อ่าว (Bay) เขื่อน (Barrage) อ่างเก็บน้ำ (Reservoir) สวนสาธารณะ (National_Park) โรงละครสัตว์ (Menagerie) ห้างสรรพสินค้า (Shopping_Center) โรงหนัง (Theatre) สถานบันเทิง (Nightlife) สวนสัตว์ (Zoo) ปราสาท (Castle) โบราณสถาน (Archaeological_Site) เกาะ (Island) ทะเลสาบ กีฬา (Stadium) วัด (Temple) ตลาด (Market) พระราชวัง (Palace) ศูนย์วัฒนธรรม (Culture_Center) อนุสาวรีย์ (Monument) และสำนักสงฆ์ (Monastery) เป็นต้น

2.2 เนมแมตซ์

การเปรียบเทียบระหว่างชื่อ 2 ชื่อเพื่อให้รู้ว่าคล้ายกัน เหมือนกัน หรือแตกต่างกันหรือไม่ นั่นอาจใช้วิธีการคำนวณค่าความเหมือน ค่าความคล้าย ค่าความแตกต่าง ซึ่งสามารถใช้ตัดสินใจ และเป็นข้อสรุปของชื่อ 2 ชื่อจะเหมือนกัน คล้ายกัน หรือแตกต่างกันหรือไม่ การใช้เนมแมตซ์ในการเปรียบเทียบชื่อ 2 ชื่ออาจทำได้ไม่ยากนัก โดยจากงานวิจัย [21] ได้ทำการแบ่งชนิดของเนมแมตซ์ไว้เป็น 4 ประเภทคือ

2.2.1 กฎพื้นฐานของตัวสะกด (Spelling Base Algorithm)

อัลกอริทึมนี้ใช้การเปรียบเทียบตัวอักษรระหว่างชื่อ 2 ชื่อ ว่ามีหลักการเขียนคล้ายกันอย่างไร อัลกอริทึมประเภทนี้ได้แก่ Guth และ Levenshtein

2.2.2 กฎพื้นฐานของการออกเสียง (Phonetic Base Algorithm)

อัลกอริทึมนี้ใช้โครงสร้างของเสียงเป็นตัวเปรียบเทียบค่าความคล้ายกัน โดยชื่อ 2 ชื่ออาจเขียนไม่เหมือนกันแต่มีการออกเสียงเหมือนกัน อัลกอริทึมประเภทนี้ได้แก่ Soundex Metaphone NYSIS และ Phonex

2.2.3 วิธีการรวมกัน (Composite Methods)

อัลกอริทึมนี้เป็นการรวมกันระหว่างกฎพื้นฐานของตัวสะกด และกฎพื้นฐานของการออกเสียง ซึ่งจะทำการเปรียบเทียบตัวอักษรระหว่างชื่อ 2 ชื่อนั้นมีหลักการเขียนคล้ายกันอย่างไรกับการใช้โครงสร้างของเสียงเป็นตัวเปรียบเทียบค่าความคล้าย โดยอัลกอริทึมนี้ได้แก้ปัญหาชื่อที่เขียนเหมือนกันแต่ออกเสียงต่างกันและชื่อที่เขียนต่างกันแต่ออกเสียงเหมือนกัน อัลกอริทึม

ประเภทนี้ได้แก่ SIMPLEX และ ISG (Index of Similarity Group)

2.1.4 วิธีการผสม (Hybrid Approaches)

อัลกอริทึมนี้เป็นการนำอัลกอริทึมหลาย ๆ แบบมารวมกัน เพื่อให้ครอบคลุมปัญหาทั้งหมดของการเกิดคำพ้องรูปและพ้องเสียง อัลกอริทึมประเภทนี้ได้แก่ LIG (Levenshtein Index of Similarity and Guth) ซึ่งพัฒนาโดย [22] ซึ่งได้นำความน่าจะเป็นในการวัดค่าความคล้ายกันของชื่อทำให้ค่าความถูกต้องในการหาชื่อคล้ายกันหรือเหมือนกันมีประสิทธิภาพมากขึ้น โดยในปัจจุบันได้มีพัฒนาเป็น LIG3 และสามารถใช้ได้กับชื่อหลายประเทศ

ในการวิจัยครั้งนี้จะใช้วิธีการรวมกันจากกฎพื้นฐานของตัวสะกดและกฎพื้นฐานของการออกเสียงในการสร้างคำที่สามารถเขียนได้หลากหลายทั้งภาษาไทยและภาษาอังกฤษ เพื่อแก้ปัญหาการพิมพ์ชื่อผิดหรือออกเสียงผิด โดยมีหลักการคือนักท่องเที่ยวป้อนคำค้นหา ระบบจะทำการสร้างคำต่าง ๆ ที่สามารถเขียนได้หลายรูปแบบ เพื่อจะนำไปแก้ปัญหาเว็บไซต์ที่มีการเขียนหรือสะกดผิดและไม่สามารถสืบค้นเจอให้ได้ผลลัพธ์ที่ไม่ผิดพลาด ตรงกับความต้องการของผู้ใช้ เช่น พิมพ์ พิธโลก หรือ Phisanulok จะทำให้รู้ได้ว่าต้องการค้นหาจังหวัดพิษณุโลก หรือ Phitsanulok เป็นต้น ดังแสดงตารางที่ 1 และตารางที่ 2

จากตารางที่ 1 และตารางที่ 2 แสดงตัวอย่างการค้นหาจังหวัดพิษณุโลกในกูเกิ้ลที่มีการเขียนหลากหลายแบบตามหลักการออกเสียงในภาษาอังกฤษและภาษาไทย ซึ่งพบว่าบางรายการที่เขียนไม่ถูกต้องก็ยังพบจำนวนรายการที่แสดงข้อมูลจังหวัดพิษณุโลก ซึ่งถ้าค้นหาสถานที่เกี่ยวกับพิษณุโลกนั้นอาจทำให้ต้องเสียเวลาค้นหาหรือค้นหาไม่เจอก็เป็นได้ ผู้วิจัยจึงพัฒนาอัลกอริทึมใหม่ในการสร้างคำที่สามารถเขียนได้หลายรูปแบบ เช่น พ้องรูปหรือพ้องเสียงเพื่อที่จะทำให้สามารถค้นข้อมูลเกี่ยวกับคำ ๆ นี้เจอในกูเกิ้ลได้ในการค้นหาเพียงครั้งเดียวเท่านั้น ซึ่งจะทำให้การค้นหามีประสิทธิภาพสูงสุด

ตารางที่ 1. แสดงตัวอย่างจำนวนรายการที่ค้นพบโดยกุ้เกิ้ลของคำค้นหา พิศณุโลก (ภาษาไทย)

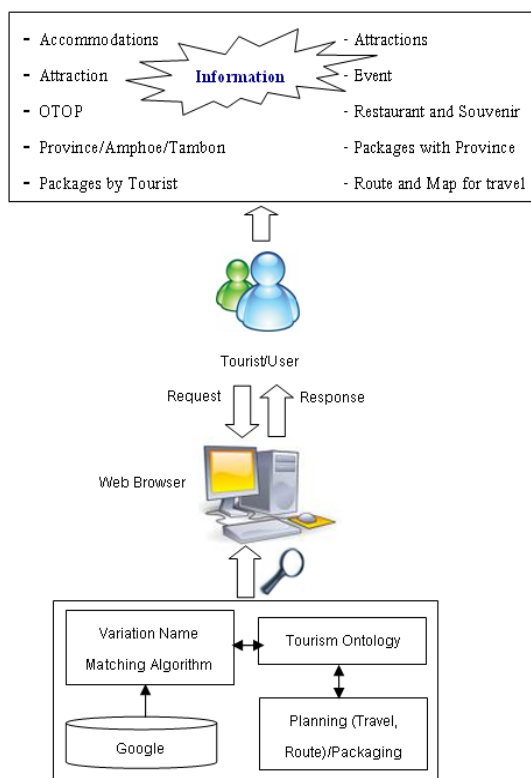
ลำดับ	ชื่อคล้ายคลึง	จำนวนรายการที่พบ (กูเกิ้ล, เมษายน 2553)
1	พิชชะณุโลก	10,100,000
2	พิศโลก	5,770,000
3	พิชญโลก	5,470,000
4	พิศชะณุโลก	2,290,000
5	พิชชะณุโลก	2,200,000
6	พิสชะณุโลก	1,230,000
7	พิชญโลก	931,000
8	พิชะณุโลก	110,000
9	พิศชะณุโลก	106,000
10	พิคชะณุโลก	76,200

ตารางที่ 2. แสดงตัวอย่างจำนวนรายการที่ค้นพบโดยกู้เกิ้ลของคำค้นหา Phitsanulok (ภาษาอังกฤษ)

ลำดับ	ชื่อคล้ายคลึง	จำนวนรายการที่พบ (กูเกิ้ล, เมษายน 2553)
1	phitsanulok	1,240,000
2	phisanulok	603,000
3	pitsanulok	117,000
4	pisanulok	17,200
5	phitsanurok	1,910
6	phisanuklok	105
7	phitzanulok	53
8	phitsanoolok	6
9	phitsanoklok	3
10	phisanooklok	1

2.3 สถาปัตยกรรมระบบ

จากรูปที่ 5 แสดงสถาปัตยกรรมของระบบ เมื่อผู้ใช้งานต้องการข้อมูลแหล่งท่องเที่ยวในประเทศไทย ผู้ใช้จะป้อนคำค้นหาจากนั้นระบบจะดึงความรู้และข้อมูลต่าง ๆ ที่ถูกเก็บไว้ในออนโทโลยีท่องเที่ยว โดยผลลัพธ์ที่ได้นั้นจะถูกนำไปค้นหาผลลัพธ์ที่มีการเขียนหรือสะกดผิดในเว็บไซต์ต่าง ๆ โดยใช้ variation name matching algorithm มาทำการหาคำที่สามารถเขียนผิดได้หลากหลายรูปแบบในเว็บไซต์ต่างๆ และทำการดึงข้อมูลต่าง ๆ เหล่านั้นจากกู้เกิ้ลแล้วส่งต่อไปให้กับผู้ใช้



รูปที่ 5. แสดงสถาปัตยกรรมของระบบ

ระบบนี้ยังสามารถนำเอาข้อมูลต่าง ๆ ไปใช้ในการวางแผนการเดินทางท่องเที่ยวและระบบสามารถจัดแพ็คเกจการท่องเที่ยวเพื่อตอบสนองความต้องการของนักท่องเที่ยวภายในจังหวัดนั้นได้ตรงตามความต้องการมากที่สุดอีกด้วย

หลักการทำงานของอัลกอริทึมที่ใช้ในการค้นหาข้อมูลที่ถูกเก็บไว้ในออนโทโลยีท่องเที่ยว มีดังต่อไปนี้

1. ป้อนชื่อจังหวัดที่ต้องการค้นหา
2. รับชื่อจังหวัดไปตรวจสอบชื่อที่มีความคล้ายคลึงกันและค้นพบในกู้เกิ้ลด้วยอัลกอริทึม variation name matching

3.แสดงผลเป็นจำนวนรายการของชื่อจังหวัดและข้อมูลในแต่ละรายการที่ค้นหาเจอในข้อที่ 2

4.แสดงข้อมูลพื้นฐานของจังหวัดที่ค้นหาในข้อที่ 1 เช่น สถานที่ท่องเที่ยว ที่พัก ร้านอาหาร เป็นต้น จากข้อมูลออนโทโลยีท่องเที่ยว (tourism ontology)

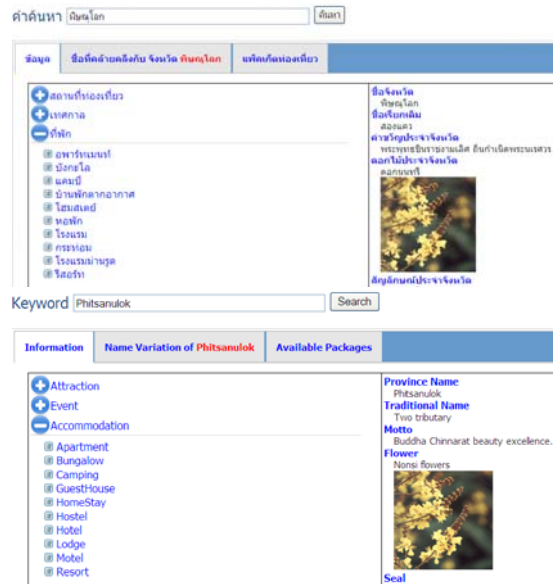
5.แสดงข้อมูลและแผนการท่องเที่ยวที่เป็นไปได้ในการท่องเที่ยวภายในจังหวัดนั้น เมื่อมีการกดปุ่มค้นหาแพ็คเกจภายในจังหวัดที่ค้นหา

3. การทดสอบและผลลัพธ์

ในหัวข้อนี้จะแสดงผลลัพธ์ของระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยด้วยหลักการออนโทโลยีและเนมแมทชิงซึ่งพัฒนาด้วยภาษา PHP โดยใช้ Protégé และ MySQL ในการเก็บข้อมูลท่องเที่ยวต่าง ๆ

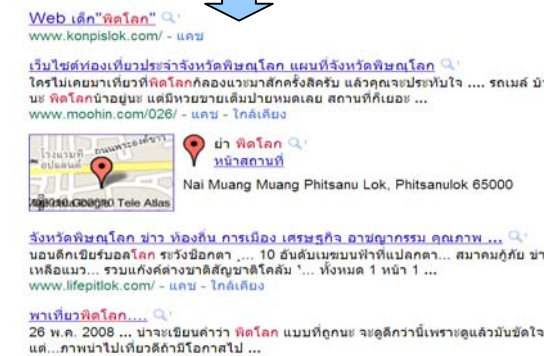
รูปที่ 6 แสดงผลลัพธ์ของการค้นหาเมื่อผู้ใช้ป้อนคำว่า “พิษณุโลก” (ภาษาไทย) หรือ “Phitsanulok” (ภาษาอังกฤษ) ระบบจะแสดงข้อมูลท่องเที่ยวที่เกี่ยวข้องกับจังหวัดนั้นๆ เช่น สถานที่ท่องเที่ยวต่างๆ (Attraction) เทศกาล (Event) โรงแรมที่พัก (Accommodation) เป็นต้น อีกทั้งระบบยังแสดงข้อมูลเบื้องต้นเกี่ยวกับจังหวัดพิษณุโลกอีกด้วย อาทิเช่น ชื่อเรียกเดิมของจังหวัด (Traditional Name) คำขวัญประจำจังหวัด (Motto) ดอกไม้ประจำจังหวัด (Flower) ต้นไม้ประจำจังหวัด (Tree) และ สัญลักษณ์หรือตราประจำจังหวัด (Seal) เป็นต้น

นอกจากนี้ผู้ใช้งานยังสามารถเรียกดูข้อมูลท่องเที่ยวจากชื่อของจังหวัดที่มีการเขียนผิดหรือสะกดผิดแต่พบข้อมูลเหล่านั้นในกูเกิ้ล โดยเลือกกดปุ่ม “ชื่อที่คล้ายคลึงกับ จังหวัด พิษณุโลก” ระบบจะแสดงข้อมูลชื่อที่มีการเขียนได้หลากหลายตามหลักภาษาไทยหรืออังกฤษ เช่น พิษณุโลก (Phitsanulok), พิษณุโรก (Phitsanurok), พิดโลก (Phitlok) เป็นต้น โดยระบบจะสร้างชื่อหลากหลายแบบอัตโนมัติด้วยอัลกอริทึม variation name matching พร้อมทั้งแสดงจำนวนรายการที่พบจากกูเกิ้ล ซึ่งสามารถเลือกดูข้อมูลในแต่ละรายการข้อมูลที่มีชื่อคล้ายคลึงกับ พิษณุโลก (ภาษาไทย) หรือ Phitsanulok (ภาษาอังกฤษ) ดังแสดงในรูปที่ 7 และ 8 ซึ่งทำให้เกิดความสะดวกและเป็นประโยชน์สำหรับนักท่องเที่ยวชาวต่างประเทศมากยิ่งขึ้น



รูปที่ 6. แสดงหน้าจอของระบบค้นหาทั้งภาษาไทยและภาษาอังกฤษ

ลำดับ	ชื่อที่คล้ายคลึงกับ จังหวัด	จำนวนรายการที่พบจากกูเกิ้ล
1	พิษณุโลก	10,900,000 รายการ
2	พิษณุโลก	8,760,000 รายการ
3	พิษณุโลก	8,350,000 รายการ
4	พิษณุโลก	7,900,000 รายการ
5	พิษณุโลก	6,810,000 รายการ
6	พิสโลก	4,690,000 รายการ
7	พิษณุโลก	1,680,000 รายการ
8	พิษณุโลก	1,730,000 รายการ
9	พิษณุโลก	897,000 รายการ
10	พิษณุโลก	91,300 รายการ
11	พิษณุโลก	68,000 รายการ



รูปที่ 7. แสดงหน้าจอชื่อที่คล้ายคลึงกับ พิษณุโลก

Name Variation of		
No.	Name Variation	Records were from Google
1	phitsanulok	5,110,000 records
2	phisarnulok	186,000 records
3	phitsanulok	32,800 records
4	phisarnulok	5,510 records
5	phitsanulok	1,400 records
6	phisarnulok	134 records
7	phisarnulok	5 records
8	phisarnulok	5 records
9	phisarnulok	4 records



King Tour ไปตามทาง ท้ายากเดิน - Phitsanulok

Photo Album, Phitsanulok, Mar 15, '07 1:08 AM for everyone. Slideshow. ล่องแก่ง ลำน้ำเข็ก ศึกษาธรรมชาติ วน. อุทยานแห่งชาติ ขากลับแวนบัสการพระพรหมธรรมา ... kinginlove.multiply.com/photos/album/9/Phitsanulok - แดง

Phitsanulok Hotel Thailand - Discount Hotels Phitsanulok Travel Guide

Special Rates at Phitsanulok Hotels, Thailand. Get Discount Rates at Asia-Discovery.com with Online Reservation. www.asia-discovery.com/Thailand/phitsanulok/ - แดง - โกลด์เดียม

Pol Phitsanulok | Facebook

เพื่อน: จักรกฤษณ์ ศักดิ์ภูงา, Chavasan Kiattinat, BoNe NoNg LiNg, Benjamas Kanpanaphai

Pol Phitsanulok ใช้ Facebook เข้าร่วม Facebook เพื่อติดต่อกับ Pol Phitsanulok และ คนอื่นๆ ที่คุณอาจจะรู้จัก รายละเอียดรายการหาบนเตาแก็ก Pol Phitsanulok has ... th-th.facebook.com/people/Pol-Phitsanulok/100000791483667 - แดง

รูปที่ 8. แสดงหน้าจอชื่อที่คล้ายคลึงกับ Phitsanulok

รูปที่ 7 และ 8 แสดงหน้าจอรายการชื่อที่คล้ายคลึงกับ จังหวัดพิษณุโลก หรือ phitsanulok ในภาษาไทยและภาษาอังกฤษ ซึ่งพบว่าบางรายการที่เขียนคล้ายคลึงกับจังหวัดพิษณุโลก (phitsanulok) เช่น พิธโลก (phitlok) เป็นคำที่สะกดผิดแต่เมื่อเข้าไปเรียกดูข้อมูลภายในลิงค์แล้ว ก็ยังพบข้อมูลเกี่ยวกับการท่องเที่ยวของจังหวัดพิษณุโลกอยู่ด้วยซึ่งเราสามารถสรุปได้ว่าจังหวัดพิษณุโลก มีการเรียกสั้นได้อีกว่าจังหวัด พิธโลก เป็นต้น แต่อย่างไรก็ตามจำนวนรายการค้นหาที่ได้จากกูเกิ้ลนั้นยังมีบางรายการชื่อบางชื่อเช่น พิษณุโรก ที่เมื่อเข้าไปดูภายในลิงค์แล้ว กลับไม่พบข้อมูลไม่เกี่ยวข้องกันกับพิษณุโลก หรือการท่องเที่ยวปะปนอยู่เลย นั่นเป็นเพราะระบบกูเกิ้ลไม่มีการกรองข้อมูลที่สั้นได้ให้มีความถูกต้องนั่นเอง

สำหรับระบบจัดแพ็คเกจท่องเที่ยวขึ้นเป็นเพียงระบบต้นแบบที่เก็บข้อมูลแพ็คเกจท่องเที่ยวต่าง ๆ ภายในจังหวัดในประเทศไทยที่มีนักท่องเที่ยวนิยมมาใช้บริการหรือมาเที่ยวมากที่สุด มีไว้เป็นสถิติว่ามาจังหวัดนี้แล้วควรจะทำเที่ยวที่ไหนหรือแบบไหนพร้อมแนะนำเส้นทางโดยใช้แผนที่จากกูเกิ้ลแสดงลูกศรลำดับการเดินทาง เพื่อเป็นข้อมูลให้นักท่องเที่ยวสามารถตัดสินใจท่องเที่ยวตามสถานที่นั้น ๆ ได้อย่างสะดวกสบาย ดังแสดงในรูปที่ 9



คำค้นหา พิษณุโลก ค้นหา

ข้อมูล ชื่อที่คล้ายคลึงกับ จังหวัด พิษณุโลก แพ็คเกจท่องเที่ยว

แพ็คเกจในจังหวัด พิษณุโลก

← อุทิมร้อมถ้ำ → อุทิมร้อม → เขาค้อ → อุทยานหลวง → อุทยาน → เขาค้อ → เขาค้อ → เขาค้อ (3 วัน 3 คืน) [แสดงรายการแพ็คเกจ](#)

← ไท้พระพุทธรูปเขาค้อ → ล่องแก่งลำน้ำเข็ก → เขาค้อ → เขาค้อ → เขาค้อ → เขาค้อ (2 วัน 1 คืน) [แสดงรายการแพ็คเกจ](#)

← ล่องแก่งลำน้ำเข็ก → เขาค้อ → เขาค้อ → เขาค้อ → เขาค้อ (2 วัน 1 คืน) [แสดงรายการแพ็คเกจ](#)

Keyword Phitsanulok Search

Information Name Variation of Phitsanulok Available Packages

Travel packages in phitsanulok

← Phu Hin Rong Kla → Phu Thap Bock → Khao Kho → Thung Salang Luang → Thung Nang → Route Waterfall → Buddha Chinnarat (3 days 3 nights) [Travel Route and Map](#)

← Buddha Chinnarat → River Rafting Khao Kho → Khao Kho → Jethi Kan Karnchanapisek Khao Kho → Jethi Kan Karnchanapisek → International Library → Commemorate the sacrifice Khao Kho (2 days 1 night) [Travel Route and Map](#)

← Khao Kho River Cruises → Phitsanulok → Phu Hin Rong Kla → Phu Thap Bock (2 days 1 night) [Travel Route and Map](#)

รูปที่ 9. แสดงหน้าจอแพ็คเกจท่องเที่ยว (ภาษาไทยและภาษาอังกฤษ)

4. บทสรุป

ในงานวิจัยนี้ผู้วิจัยได้นำเสนอผลของการพัฒนาระบบค้นหาสถานที่ท่องเที่ยวในประเทศไทยด้วยหลักการออนไลน์และเนมเมทซิง โดยมีวัตถุประสงค์เพื่อใช้ในการค้นหาแหล่งท่องเที่ยวต่าง ๆ ได้ครบถ้วนสมบูรณ์และช่วยให้เว็บไซต์ต่าง ๆ ที่มีการเขียนผิดสามารถถูกค้นเจอในครั้งเดียว โดยไม่ต้องเสียเวลาสืบค้นหลาย ๆ ครั้งตามเสิร์ชเอนจิน ต่าง ๆ ซึ่งประสิทธิภาพของงานวิจัยนี้ประเมินได้จากการพยายามทำให้ข้อมูลต่าง ๆ ที่ได้จากการค้นหาชื่อจังหวัดที่อาจจะมีการเขียนหรือสะกดคล้ายคลึงกันหรือที่เกี่ยวข้องสามารถค้นหาเจอได้ครบถ้วนมากที่สุด อีกทั้งข้อมูลพื้นฐานของจังหวัดนั้น ๆ จะถูกแสดงอย่างมีความสัมพันธ์กันอย่างถูกต้องครบถ้วน อาทิเช่น คำขวัญประจำจังหวัด สัญลักษณ์ประจำจังหวัด ต้นไม้ประจำจังหวัด เป็นต้น

แต่อย่างไรก็ตามระบบการจัดแพ็คเกจในงานวิจัยนี้เป็น การนำเอาข้อมูลมาจัดเรียงตามความนิยมของจังหวัดนั้น ๆ ซึ่งเป็นแบบสถิติเท่านั้น ในอนาคตผู้วิจัยจะทำการพัฒนาระบบ โดยเพิ่มในส่วนของการจัดแพ็คเกจให้เป็นแบบ

ไดนามิกเพื่อให้ได้ข้อมูลที่สอดคล้องและตรงกับความต้องการของนักท่องเที่ยวมากที่สุดและเพื่อลดเวลาในการค้นหาข้อมูลในการตัดสินใจของนักท่องเที่ยวและพัฒนากระบวนการกรองข้อมูลที่ได้จากการค้นหาในกูเกิ้ลโดยตัดรายการที่ไม่เกี่ยวข้องกับการท่องเที่ยวออกซึ่งจะได้เป็นจำนวนรายการที่ค้นพบอย่างแท้จริง รวมถึงจะวัดประสิทธิภาพของอัลกอริทึม variation name matching โดยใช้หลักการ F-measure อีกด้วย

เอกสารอ้างอิง

- [1] สุวัฒน์ จุฑากรณ์และจริญญา เจริญสุกใส, “เอกสารการสอนชุดวิชา ความรู้เบื้องต้นเกี่ยวกับอุตสาหกรรมท่องเที่ยว”, 2550
- [2] การท่องเที่ยวแห่งประเทศไทย, “โครงการเที่ยวไทยคลิกกรีน เศรษฐกิจไทยคลิก”, สืบค้นเมื่อ 20 ธันวาคม 2552, จาก <http://thai.tourismthailand.org>
- [3] C. Snae, M. Brückner, and P. Wongthongtham, “Local organization and business ontology (LOBO)”, *IEEE International Conference on Digital Ecosystems and Technologies*, Rydges Tradewinds Cairns, Australia, Feb. 2007, pp. 292-295.
- [4] C. Snae, N. Singhadech, B. Emapana, and M. Brückner, “Interactive transliteration tools for explanation level language system (IT-TELLS)”, *International Technical Conference on Circuits/Systems Computers and Communication*, Chiang Mai, Thailand, Jul. 2006, pp. 245-248.
- [5] C. Snae and M. Brückner, “LOWCOST: Local organization search with consolidated ontologies for name, space and time”, *International Conference on Software Engineering*, Innsbruck, Austria, Feb. 2007, pp. 13-15
- [6] M. Brückner, C. Snae and J. Payakpate, “Ontology-based name matching of toponyms for geographical information systems (ONTO-GIS)”, *LADIS International Conference ICT, Society and Human Beings*, Amsterdam, The Netherlands, Jul. 2008, pp. 108-114
- [7] จุฬามาศ ศิริรัชนิกรและฤทธิกร เล้าอรุณ, “ระบบการค้นหาข้อมูลแบบซับซ้อนสำหรับข้อมูลการท่องเที่ยวในประเทศไทย”, สืบค้นเมื่อ 20 ธันวาคม 2552, จาก <http://cpe.kmutt.ac.th/previousproject/2006/26/Index.htm>
- [8] N. Guarino, “Formal ontology and information systems”, *the First International Conference (FOIS'98)*, Amsterdam: IOS Press, Jun.1998, pp 3-15.
- [9] H. Andrade and J. Saltz, “Towards a knowledge base management system KBMS: an ontology-aware database management system DBMS”, *the 14th Brazilian Symposium on Database*, Florianopolis, Brazil, Oct. 1999, pp 27-39.
- [10] H. Andrade and J. Saltz, “Query optimization in kess an ontology-based KBMS”, *the 15th Brazilian Symposium on Databases (SBBD'2000)*. Joao Pessoa, Brazil 2000, pp 35-48.
- [11] T. Gruber, “Ontology”, Retrieved October, 2 2009, from <http://tomgruber.org/writing/ontology-definition-2007.htm>
- [12] B. Chandrasekaran, J.R. Josephson and V.R. Benjamins, “What are ontologies and why do we need them”, *IEEE Intelligent System*. Volume 14 no.1, pp. 20-26, 1999.
- [13] B. Smith and C. Welty, “Formal ontology in information systems”, *the 2nd International Conference on formal Ontology in Information Systems*, Ogunquit, Maine, New York: ACM Press, Oct. 2001.
- [14] N.F. Noy and D.L. McGuinness, “Ontology development: a guide to creating your first ontology”, California : Stanford Knowledge Systems Laboratory Technical Report, 2001
- [15] I. Murua, E. Llado and B. Llodra, “The semantic web for improving dynamic tourist packages commercialisation”, 2006, Retrieved December, 20 2009, from http://www.ibit.org/dades/doc/1108_ca.pdf
- [16] W. V. Siricharoen, “Learning semantic web from e-tourism” in the 2nd KES International conference on Agent and Multi-Agent Systems: Technologies and Applications, Inha University, Korea, Mar. 2008, pp. 516-525.

- [17] D. Buhalisa and R. Law, "Progress in information technology and tourism management: 20 years on and 10 years after the Internet the state of Tourism research", Tourism Management, Elsevier Ltd., 2008
- [18] T. Dema, "eTourPlan: A knowledge-based tourist route and activity planner", Master thesis, M.Sc., University of New Brunswick, 2008
- [19] M. Ball and B. Craig, "Object oriented jDREW", Retrieved January 21, 2010, from <http://jdrew.org/ooidrew>.
- [20] Protégé, "The Protégé Ontology Editor and Knowledge Acquisition System", Retrieved January, 19 2010, from <http://protege.stanford.edu>
- [21] C. Snae and M. Brückner, "Concept and rule based naming system", The Information Universe. *Journal of Issues in Informing Science and Information Technology*, Volume 3, pp. 619-634, 2006.
- [22] C. Snae, "A comparison and analysis of name matching algorithms", *International Journal of Applied Science. Engineering and Technology*, Volume 4 no. 1, pp. 252-257, 2007.