การพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

ONTOLOGY DEVELOPMENT FOR STOCK PHOTOGRAGHY BUSINESS

สุชานัน อ้นถาวร

SUCHANAN AONTHAVORN

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยศรีปทุม
พ.ศ. 2560

ลิขสิทธิ์ของมหาวิทยาลัยศรีปทุม

การพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

สุชานัน อ้นถาวร

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยศรีปทุม
พ.ศ. 2560

ลิขสิทธิ์ของมหาวิทยาลัยศรีปทุม

ONTOLOGY DEVELOPMENT FOR STOCK PHOTOGRAGHY BUSINESS

SUCHANAN AONTHAVORN

A THEMATIC SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

SCHOOL OF INFORMATION TECHNOLOGY

SRIPATUM UNIVERSITY

2017

COPYRIGHT OF SRIPATUM UNIVERSITY

หัวข้อสารนิพนธ์ การพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์
คำสำคัญ ออนโทโลยี, การแยกประเภทของภาพ, OWL
นักศึกษา สุชานัน อ้นถาวร
อาจารย์ที่ปรึกษาสารนิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.สุรศักดิ์ มังสิงห์
หลักสูตร วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะ เทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม
พ.ศ. 2560

บทคัดย่อ

บทความวิจัยนี้มีวัตถุประสงค์ที่จะพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์ ซึ่งเป็นการสืบค้นรูปภาพเชิงความหมายให้ได้ตรงตามความต้องการของผู้ใช้งานและพัฒนาออนโทโลยีสำหรับการสืบค้นรูปภาพเชิงความหมายให้สอดคล้องกับวิธีการสืบค้นรูปภาพเชิงความหมาย โดยการสืบค้นรูปภาพบนเว็บไซต์ที่ใช้หลักการสืบค้นแบบใช้คำหลักจะไม่พิจารณาถึงความหมายของรูปภาพ ทำให้ไม่สามารถอธิบายข้อมูลที่ชี้เฉพาะรูปภาพได้ อาจทำให้ผู้ใช้ได้รับผลลัพธ์ที่มากจนเกินไป และไม่ถูกต้องตามความต้องการเท่าที่ควร โดยการแก้ปัญหาดังกล่าวสามารถทำได้โดยการใช้หลักการสืบค้นเชิงความหมาย จะทำให้ผลลัพธ์ที่ได้จากการสืบค้นมีความถูกต้องมากยิ่งขึ้น ซึ่งงานวิจัยนี้ได้ดำเนินการออกแบบโครงสร้างข้อมูลออนโทโลยีสำหรับอธิบายคุณลักษณะหรือความหมายของรูปภาพ แล้วจัดเก็บคำบรรยายให้อยู่ในรูปแบบของภาษา OWL หลังจากนั้นจึงออกแบบและพัฒนาออนโทโลยีสำหรับการสืบค้นรูปภาพเชิงความหมายขึ้นมา ในการพัฒนาออนโทโลยีสำหรับการสืบค้นรูปภาพเชิงความหมายครั้งนี้ มีการทดลองการค้นหารูปภาพจากฐานข้อมูลที่จัดเก็บให้อยู่ในรูปแบบของ OWL โดยทำการสืบค้นด้วยการใช้คำหลักที่แตกต่างกันเพื่อทดลองความแม่นยำ นอกจากนี้ได้ทำการตรวจความตรงเชิงเนื้อหาจากผู้ทรงคุณวุฒิจำนวน 5 คน ผลการประเมินได้ค่า IOC = 1 แสดงว่าองค์ความรู้ในการแยกประเภทของภาพมีคุณภาพ ซึ่งเนื้อหาดังกล่าวถูกแปลงเป็นออนโทโลยีด้วยโปรแกรม Protégé (Protégé Project, 2016) ผลการวัดประสิทธิภาพในการสืบค้นด้วยวิธี Recall เท่ากับ 0.85 และวิธี Precision เท่ากับ 0.82 ผลการทดลองที่ได้คือประสิทธิภาพความถูกต้องของระบบอยู่ที่ระดับระดับดีมาก และประสิทธิภาพความแม่นยำของระบบอยู่ที่ระดับระดับดี

THEMATIC TITLE	ONTOLOGY DEVELOPMENT FOR STOCK PHOTOGRAGHY BUSINESS
KEYWORD	ONTOLOGY, CLASSIFICATION OF IMAGES, OWL
STUDENT	SUCHANAN AONTHAVORN
ADVISOR	ASST. PROF. DR. SURASAK MUNGSING
LEVEL OF STUDY	MASTER OF SCIENCE IN INFORMATION TECHNOLOGY
FACULTY	SCHOOL OF INFORMATION TECHNOLOGY SRIPATHUM UNIVERSITY
YEAR	2017

ABSTRACT

This paper aims to develop ontology for semantic image searching to meet the needs of users and to comply with semantic image searching method. Searching images on the website using the search keywords will not take into consideration the meaning of the image, which is unable to describe the specific image and may provide users with too many results and not meet the user's needs. The solution of this problem can be achieved through the use of semantic search to make the results of the search even more accurate. This research was conducted to design ontology for description or definition of the image, store narration in the form of OWL language, and then design and develop ontology based system for image search. For this research, ontology was developed as semantic technology for image search. The experiments for image search from the database, using keywords with different experimental precision, were conducted and stored results in the form of OWL. In addition, the content validity was examined by five qualified people. The result showed that IOC = 1, which means that knowledge can distinguish the type of image quality. This content was converted to an ontology using Protégé program. **(**Protégé Project, 2016) The performance in queries with Recall is 0.85 and Precision is 0.82. The result is an effective accuracy of the system is at a very good level. Precision performance of the system is at a good level.

กิตติกรรมประกาศ

สารนิพนธ์ฉบับนี้สำเร็จลงได้ด้วยความอนุเคราะห์อย่างดียิ่ง จากผู้ช่วยศาสตราจารย์ ดร.สรุศักดิ์ มังสิงห์ ซึ่งเป็นอาจารย์ที่ปรึกษา ที่ได้ให้ความกรุณาแนะนำความรู้ คำปรึกษาปัญหา แนวทางแก้ไข และคอยติดตามความคืบหน้าอย่างต่อเนื่อง และขอบพระคุณอาจารย์ทุก ๆ ท่าน ในสาขาวิชาเทคโนโลยีสารสนเทศ ที่ให้วิชาความรู้ต่าง ๆ ส่งผลให้ผู้ศึกษาได้นำความรู้มาใช้ในการจัดทำสารนิพนธ์ฉบับนี้

ขอขอบพระคุณครูบาอาจารย์ที่ประสาทวิชาความรู้แก่ข้าพเจ้าจนสามารถประสบความสำเร็จในการทำสารนิพนธ์นี้

ขอขอบพระคุณครอบครัว คุณพ่อคุณแม่ที่อบรมสั่งสอนให้ซื่อสัตย์และใฝ่เรียนใฝ่รู้ และเป็นกำลังใจ เป็นแรงผลักดันที่สำคัญยิ่ง

ขอขอบคุณเพื่อน หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ รุ่น 20 สำหรับมิตรภาพ และความช่วยเหลือซึ่งกันและกันจนจบหลักสูตร

สุชานัน อ้นถาวร

สารบัญ

หน้า

บทคัดย่อภาษาไทย I
บทคัดย่อภาษาอังกฤษ II
กิตติกรรมประกาศ..... III
สารบัญ..... IV
สารบัญตาราง VI
สารบัญภาพ VII

บทที่

1 บทนำ 1
ความเป็นมาและความสำคัญ..... 1
วัตถุประสงค์..... 2
ขอบเขตการศึกษา..... 2
ประโยชน์ที่คาดว่าจะได้รับ..... 2

2 แนวคิด ทฤษฎี เทคโนโลยี และระบบงานที่เกี่ยวข้อง..... 3
แนวคิด..... 3
ทฤษฎีที่เกี่ยวข้อง..... 4
เทคโนโลยีที่เกี่ยวข้อง..... 22
ผลการศึกษาที่เกี่ยวข้อง..... 44

3 วิธีการดำเนินการศึกษา..... 50
การดำเนินการศึกษา..... 50
เครื่องมือและอุปกรณ์ที่ใช้..... 50
วิธีดำเนินการ..... 51
ระยะเวลาในการดำเนินการศึกษา..... 65

สารบัญ (ต่อ)

บทที่		หน้า
4	ผลการศึกษา	68
	การทดลองและข้อกำหนดเบื้องต้นของการทดลอง	68
	การประเมินผล	71
5	สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	74
	สรุปและอภิปรายผล	74
	ปัญหาอุปสรรค	75
	ข้อเสนอแนะ	75
บรรณานุกรม		77
ภาคผนวก		79
	ภาคผนวก ก	80
	ภาคผนวก ข	88
ประวัติผู้วิจัย		90

สารบัญตาราง

ตารางที่		หน้า
2-1	ตารางผลจากการสืบค้นในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี.....	30
3-1	ตารางตัวอย่างพจนานุกรมคำศัพท์ จากการจำแนกประเภทภาพและ..... กำหนดคำศัพท์ (Vocabulary)	51
3-2	ตัวอย่างผลลัพธ์ของการเลือกกลุ่มคำสืบค้น.....	56
3-3	ตัวอย่างผลลัพธ์ของการสืบค้นข้อมูลรูปภาพจากเอกสาร OWL.....	57
3-4	ตัวอย่างข้อมูลที่มีความสัมพันธ์กันจากโครงสร้างการอธิบายข้อมูล OWL	58
3-5	ตัวอย่างผลลัพธ์ที่ได้จากการคิวรีข้อมูลด้วยภาษา SPARQL.....	62
3-6	แผนการดำเนินงาน.....	65
4-1	คำและวลีสืบค้นที่ใช้การทดลองออนโทโลยีสำหรับธุรกิจขายภาพ..... ออนไลน์.....	68
4-2	แสดงเกณฑ์การแปลความหมายข้อมูลและพิจารณา จากค่า Precision..... และ Recall ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์.....	70
4-3	แสดงค่าความตรงตามเนื้อหา (IOC) ของข้อมูล.....	72

สารบัญภาพ

ภาพประกอบที่ หน้า

2-1 แนวทางการศึกษาความจริงด้วยออนโทโลยี..... 5
2-2 ความสัมพันธ์ของออนโทโลยีกับการประมวลผลด้วยคอมพิวเตอร์..... 6
2-3 ภาพแสดง KB life cycle..... 10
2-4 ขั้นตอนการพัฒนาออนโทโลยี..... 11
2-5 แสดงตัวอย่างของสถาปัตยกรรมฐานความรู้…..... 13
2-6 ภาพแสดง Conceptual graph…….... 15
2-7 ภาพแสดง Predicate calculus..... 15
2-8 ตัวอย่างเดนโดรแกรมจากการจัดกลุ่มแบบลำดับชั้น..... 18
2-9 สเปคตรัมของความรู้ (Knowledge spectrum)..... 19
2-10 การจัดการความรู้เชิงความหมายสำหรับเทคโนโลยีเว็บยุคหน้า..... 20
2-11 ภาพแสดง Metadata ของ Ontology..... 22
2-12 ทิศทางการพัฒนาเทคโนโลยีเว็บ..... 23
2-13 การพัฒนาฐานความรู้ด้านการดูแลรักษาโรคเบาหวานโดยใช้เครื่องมือ... 25
สนับสนุนการพัฒนาออนโทโลยี.....
2-14 การประยุกต์ใช้งานในโปรแกรมแจ้งเตือนความจำ (Reminder) 27
2-15 ลักษณะการทำงานของ Web Crawler..... 32
2-16 ลักษณะการทำงานของ Web Crawler..... 33
2-17 ตัวอย่างเว็บไซต์ที่ทำงานแบบ Web Crawler..... 33
2-18 ตัวอย่าง Web Directory ที่มี Search Engine มากมายใช้เป็นฐานข้อมูล... 34
2-19 ตัวอย่างคำสั่งใน Tag META เพื่อประกาศข้อมูลของเว็บ..... 35
2-20 ส่วนประกอบของ Triple..... 36
2-21 ตัวอย่างการกำหนด Namespace..... 39
2-22 ตัวอย่าง Ontology Header..... 39
2-23 ตัวอย่างการกำหนดคลาสและคลาสย่อย..... 40
2-24 การกำหนดความสัมพันธ์ระหว่างทรัพยากร (Resource)..... 40

สารบัญภาพ (ต่อ)

ภาพประกอบที่		หน้า
2-25	สถาปัตยกรรมทั่วไปในการสืบค้นข้อมูล.....	42
2-26	แสดงขั้นตอนการสืบค้นข้อมูล.....	48
2-27	แสดงลำดับขั้นตอนการการสืบค้นข้อมูล.....	50
3-1	แสดงการออกแบบออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์.....	54
3-2	ตัวอย่างการอธิบายข้อมูลรูปภาพด้วยภาษา OWL.....	55
3-3	ขั้นตอนการทำงานของระบบ.....	56
3-4	ตัวอย่างโครงสร้างออนโทโลยีคลาสของรูปภาพ.....	58
3-5	ตัวอย่างการบรรยายรายละเอียดข้อมูลของ OWL.....	60
3-6	ตัวอย่างการค้นหาข้อมูลรูปภาพจากซับคลาสหมวดลักษณะของรูปภาพ	61
3-7	อธิบายการค้นหาข้อมูลรูปภาพจากซับคลาสหมวดสัตว์.....	61
3-8	ภาพตัวอย่างหน้าจอรับคำค้น.....	63
3-9	กรอบแนวคิดในการวิจัยการพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพ..... ออนไลน์.....	66
3-10	กรอบแนวคิดการพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์.....	67
4-1	ตัวอย่างการสืบค้นด้วยการระบุเงื่อนไขการสืบค้น.....	71
ผ-1	หน้าจอโปรแกรม Protégé.....	87
ผ-2	การสร้างคลาสและซับคลาส.....	92
ผ-3	ลำดับชั้นของคลาสที่สร้างตามออนโทโลยีที่ออกแบบไว้.....	93

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันมีความนิยมหันมาซื้อภาพผ่านทางเว็บไซต์กันมากขึ้น เนื่องจากมีความหลากหลายในตัวเลือก สะดวกรวดเร็ว พร้อมใช้งานและประหยัดงบประมาณ ภาพเหล่านี้เรียกกันว่าภาพสต๊อก (Stock photography) และเว็บไซต์ผู้ให้บริการขายภาพสต๊อกเหล่านี้เรียกกันว่า Microstock (Shutterstock, 2014) เปรียบเสมือนพ่อค้าคนกลางที่ไม่ได้ซื้อมาแต่ได้ขายไป การขายภาพสามารถเริ่มได้จากบุคคลทั่วไป โดยมีการตรวจสอบภาพทุกภาพตามมาตรฐาน ก่อนที่จะอนุมัติและนำภาพที่ได้รับการอนุมัติไปวางขายผ่านเว็บไซต์และมีข้อมูลที่เป็นข้อมูลรูปภาพ

ข้อมูลรูปภาพเป็นข้อมูลรูปแบบหนึ่งที่ผู้ใช้งานสามารถรับรู้ข้อมูลทางตาด้วยการมองเห็น เช่น รูปภาพจากหนังสือ โทรทัศน์ อินเตอร์เน็ต ฯลฯ ปัจจุบันความก้าวหน้าของเทคโนโลยีคอมพิวเตอร์ช่วยให้ระบบคอมพิวเตอร์สามารถประมวลผล และเก็บรวบรวมข้อมูลรูปภาพจำนวนมากไว้ในฐานข้อมูลรูปภาพ เพื่อให้ผู้ใช้งานเรียกใช้งานต่อไป ในการใช้งานข้อมูลรูปภาพนั้น ผู้ใช้งานมีความจำเป็นและความต้องการสืบค้นข้อมูลรูปภาพที่ต้องการสืบค้น อาจจะเป็นการสืบค้นด้วยข้อความที่ระบุคำสำคัญ (Keyword) ที่ต้องการสืบค้น หรือเป็นการสืบค้นด้วยรายละเอียดของรูปภาพโดยผู้ใช้งานจำเป็นต้องระบุคุณสมบัติของรูปภาพ เช่น สี พื้นผิว รูปทรง ที่ปรากฏในรูปภาพ ในการสืบค้นข้อมูลรูปภาพด้วยวิธีดังกล่าว เป็นวิธีการที่ไม่เหมาะสมสำหรับการสืบค้นข้อมูลรูปภาพบนฐานข้อมูลรูปภาพที่มีขนาดใหญ่ เนื่องจากการให้คำจำกัดความรูปภาพขึ้นกับแต่ละบุคคล ซึ่งอาจจะให้คำกำจัดความที่ไม่ตรงกัน เมื่อฐานข้อมูลมีขนาดใหญ่มากขึ้นการให้คำจำกัดความรูปภาพแต่ละรูปทำได้ค่อนข้าง ยากและใช้เวลานาน

ผู้วิจัยจึงได้ทำการพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์ โดยใช้ทฤษฎีการจำแนกแบบออนโทโลยี (ธีรวิชญ์ วงษา, 2557) ออนโทโลยีคือทฤษฎีทางตรรกวิทยาที่ใช้กำหนดความหมายของการอธิบายคำศัพท์จากขอบเขตที่สนใจ ซึ่งมีการกำหนดการอธิบายความหมายนั้นด้วยรูปแบบอย่างเป็นทางการที่มีโครงสร้างและความสัมพันธ์แบบลำดับชั้น ทำให้มีความสามารถในการถ่ายทอดคุณสมบัติ โดยใช้นิยามแนวคิดให้อยู่ในรูปแบบของคลาส, ความสัมพันธ์ระหว่างคลาส, คุณสมบัติของคลาส และสล๊อต (Stamford, 2014) ทำให้ช่วยให้พัฒนาการสืบค้นรูปภาพให้มีประสิทธิภาพมากยิ่งขึ้นได้ผลลัพธ์จากการสืบค้นรูปภาพตรงตามความต้องการของผู้ใช้งานนำไปใช้กับการสืบค้นรูปภาพต่อไป

วัตถุประสงค์

(1) เพื่อการพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

(2) เพื่อออกแบบและพัฒนาการสืบค้นรูปภาพที่มีประสิทธิภาพ สามารถสืบค้นรูปภาพได้ตรงตามความต้องการของผู้ใช้งาน

ขอบเขตของการศึกษา

สามารถนำหลักการของออนโทโลยีที่เป็นแนวคิดที่ถูกนำมาประยุกต์ใช้เพื่อการจัดการข้อมูลต่างๆ สามารถจัดการและนำเสนอข้อมูลในรูปแบบของกลุ่มแนวคิดและความสัมพันธ์ระหว่างกันที่อยู่ภายใต้ขอบเขตที่สนใจ โดยการรวบรวมข้อมูลและสร้างความสัมพันธ์ให้อยู่ในรูปแบบลำดับชั้น

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อการออกแบบและพัฒนาระบบสำหรับจัดหมวดหมู่ของภาพในการขายภาพออนไลน์

2. ได้แนวทางใหม่ในการเข้าถึงข้อมูลเพื่อวิเคราะห์ภาพได้โดยง่ายเพื่อความสะดวกในการค้นหาและได้ภาพที่ตรงกับความต้องการที่สุด

3. มีระบบวิเคราะห์ภาพที่สามารถรอบรับความซับซ้อนของข้อมูล

บทที่ 2

แนวคิด ทฤษฎี การศึกษาที่เกี่ยวข้อง

ในการศึกษาครั้งนี้ ผู้ศึกษาได้ค้นคว้าแนวคิดและทฤษฎีที่เกี่ยวข้อง เพื่อเป็นแนวทางในการศึกษาการพัฒนาออนโทโลยีสำหรับธุรกิจการขายภาพออนไลน์ (Ontology development for stock photography business) โดยมีรายละเอียดที่เกี่ยวข้องดังนี้

แนวคิด

ปัจจุบันโลกได้ก้าวเข้าสู่ยุคของเศรษฐกิจฐานความรู้ซึ่งเป็นการใช้ความรู้ในการสร้างมูลค่าเพิ่ม เทคโนโลยี ความรู้ หรือปัญญาประดิษฐ์มีส่วนช่วยในการส่งผ่านความรู้จากคนสู่คอมพิวเตอร์และใช้ระบบสนับสนุนการตัดสินใจในการเพิ่มมูลค่าทางเศรษฐกิจ การจัดการความรู้เป็นกระบวนการของการจัดการ เพื่อรวบรวมสร้างความรู้ นำเสนอ แลกเปลี่ยน และการนำมาประยุกต์ใช้เพื่อนำไปสู่การจัดการสารสนเทศที่มีประสิทธิภาพ ถูกนำมาประยุกต์ใช้ในองค์กรเพื่อเพิ่มประสิทธิภาพในการแข่งขัน กระตุ้นการคิดค้นนวัตกรรม ใหม่ๆ แบ่งปันแลกเปลี่ยนองค์ความรู้ที่ได้รับบูรณาการและพัฒนาองค์ความรู้อย่างต่อเนื่อง การนำเสนอความรู้เป็นกระบวนการหนึ่งในการจัดการความรู้เพื่อการเผยแพร่ความรู้ที่รวบรวมมาได้จากแหล่งความรู้ต่างๆ อาจถูกนำเสนอด้วยเสียง ข้อความ หรือสัญลักษณ์ เพื่อการนำเสนอที่มีประสิทธิภาพตามความเหมาะสมในงานวิจัยที่เกี่ยวข้องกับการนำเสนอความรู้มักจะเกี่ยวข้องกับการวิเคราะห์การใช้สัญลักษณ์ในการนำเสนอกลุ่มข้อเท็จจริงในขอบเขตของความรู้อย่างไรให้มีประสิทธิภาพ

ออนโทโลยีเป็นแนวคิดที่ถูกนำมาประยุกต์ใช้เพื่อการจัดการความรู้เป็นที่นิยมใช้และศึกษาอย่างแพร่ หลาย สามารถจัดการและนำเสนอความรู้ในรูปแบบของกลุ่มแนวคิดและความสัมพันธ์ระหว่างกันที่อยู่ภาย ใต้ขอบเขตที่สนใจ โดยการรวบรวมข้อมูลและสร้างความสัมพันธ์ให้อยู่ในรูปแบบลำดับชั้น มักถูกใช้ในสาขาปัญญาประดิษฐ์ เว็บเชิงความหมาย (Semantic Web) วิศวกรรมซอฟต์แวร์และสารสนเทศทางชีวการแพทย์ ออนโทโลยีเป็นระบบคำศัพท์เชิงความหมายใช้เป็นโครงร่างพื้นฐานในการอธิบายความรู้เฉพาะด้าน โดยกลุ่มของคำที่มีโครงสร้างแบบลำดับชั้น

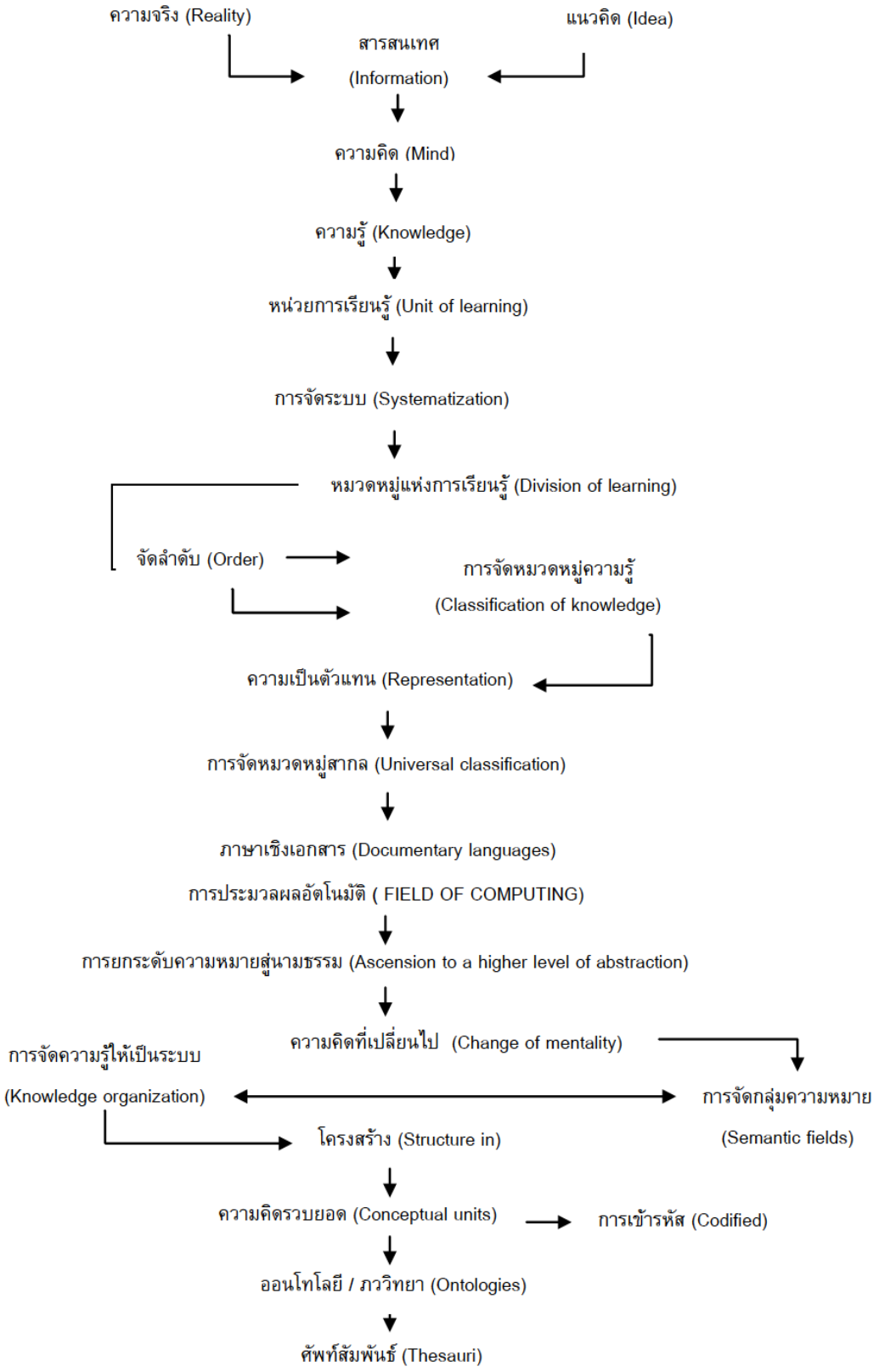
สำหรับใช้อธิบายขอบเขตเนื้อหาที่สนใจ สามารถนำมาประยุกต์กับระบบต่างๆ ได้ เช่น การจัดรวมข้อมูลสารสนเทศการจำแนกเอกสารการสกัดข้อสนเทศ เป็นต้น

ปัญหาสำคัญอย่างหนึ่งในการสร้างออนโทโลยี คือการพึ่งพาผู้เชี่ยวชาญในการสกัดองค์ความรู้และจัดประเภทความรู้ให้เป็นหมวดหมู่เพื่อให้เหมาะสมต่อการนำเอาความรู้ไปใช้ประโยชน์ การคัดแยกและจัด กลุ่มเนื้อหาที่มีความใกล้เคียงเพื่อเชื่อมโยงเนื้อหา ซึ่งบางครั้งการหาผู้เชี่ยวชาญในเรื่องใดเรื่องหนึ่งอาจเป็นเรื่องยาก หากสามารถใช้เทคโนโลยีเข้ามาช่วยในการจัดกลุ่มเนื้อหาให้เข้าใกล้ความเป็นอัตโนมัติก็จะสามารถ ลดปัญหาการพึ่งพาผู้เชี่ยวชาญไปได้ ในงานวิจัยนี้ผู้วิจัยได้ค้นคว้าและศึกษาวิธีการสร้างออนโทโลยี โดย พยายามลดบทบาทของคนลงและให้คอมพิวเตอร์เข้ามาจัดการ โดยเลือกใช้วิธีการจัดกลุ่มแบบลำดับชั้นเข้ามา ช่วยในการจัดกลุ่มข้อมูลและขึ้นโครงสร้างแบบลำดับชั้น ในรูปแบบของเดนโดรแกรมซึ่งสอดคล้องกับโครงสร้างของออนโทโลยี ซึ่งแบ่งกระบวนการทำงานเป็นขั้นตอน ได้แก่ การรวบรวมเอกสาร การตัดคำการนับความถี่ ของคำ การเลือกคำสำคัญ สร้างตารางความถี่ที่เกิดขึ้นในเอกสาร (if-idf) จัดกลุ่มด้วยวิธีการจัดกลุ่มแบบลำดับชั้นและแปลงผลลัพธ์ให้อยู่ในภาษาเว็บออนโทโลยีเพื่อสร้างเป็นออนโทโลยีที่อยู่ในรูปของภาษาเว็บออนโทโลยีต่อไป

ทฤษฎีที่เกี่ยวข้อง

ออนโทโลยี หมายถึง แนวคิดที่ใช้กำหนดความหมายที่เป็นทางการของคำศัพท์พร้อมทั้งประกาศคุณลักษณะที่ชัดแจ้งเพื่อนำมาใช้ในการอธิบายความเป็นตัวแทนของแนวคิด (Concepts) หรือแบบจำลอง (Model) ของกลุ่มชุมชนสารสนเทศที่ใช้ร่วมกัน (Information Communities) ซึ่งโครงสร้างความสัมพันธ์ดังกล่าว เครื่องคอมพิวเตอร์สามารถสามารถเข้าใจและแปลความได้โดยใช้คลาส (Class) ความสัมพันธ์ระหว่างคลาส รวมถึงลำดับชั้นของคลาสและคุณสมบัติ (Properties) ของคลาส ความรู้ที่ได้จากออนโทโลยีมีขอบเขตอยู่เฉพาะทาง (Domain) ซึ่งช่วยสนับสนุนกระบวนการค้นคืนสารสนเทศ (Information Retrieval) ในแง่ของการตัดทอนคำศัพท์ที่สับสน หรือบรรยายเชิงความหมายจากหลายแนวคิด (Concepts) ให้สอดคล้องกันภายใต้แนวคิด (Concept) เพียงหนึ่งเดียว ทั้งยังมีบทบาทสำคัญต่อการพัฒนาระบบความรู้ (Knowledge Based Systems) ในแง่ของการนำกลับมาใช้ใหม่ (Reusable) และเพิ่มเติมองค์ประกอบได้ภายหลัง ส่วนภาษาที่ใช้ในออนโทโลยีเพื่อบรรยายข้อมูลเชิงความหมาย ได้แก่ XML (Extensible Markup Language) RDF (Resource Description Framework) และ OWL (Web Ontology Language) ในการพัฒนาออนโทโลยีแต่ละประเภทจะขึ้นอยู่กับวัตถุประสงค์ ขอบเขตของความรู้ บริบทแวดล้อม และความพร้อมในการพัฒนา

ภาพประกอบที่ 2-1 แนวทางการศึกษาความจริงด้วยออนโทโลยี ปรับปรุงมาจาก Currás (2010)



ภาพประกอบที่ 2-2 ความสัมพันธ์ของออนโทโลยีกับการประมวลผลด้วยคอมพิวเตอร์ ปรับปรุงมาจาก Currás (2010)

ออนโทโลยีเป็นเทคโนโลยีทางด้านการพัฒนาภาษาเชิงความหมายสมัยใหม่ โดยเป็นภาษาที่เครื่อง คอมพิวเตอร์สามารถตีความหมายและทำตามคำสั่งได้ โดยลักษณะของออนโทโลยีคือการนิยามรูปแบบ (Model) ภายในขอบเขตขององค์ความรู้เพื่ออธิบายสิ่งที่เราสนใจ (Domain) ให้ได้ใจความและถูกต้องมากที่สุด ตัวอย่างเช่น ในการนิยามออนโทโลยีของสิ่งมีชีวิตโดยสิ่งมีชีวิตนั้นแบ่งเป็น 2 ประเภทคือสัตว์และพืช โดยจะทำการนิยามสัตว์เพิ่มขึ้นอีกคือสัตว์แบ่งเป็น 3 ประเภท คือ 1) สัตว์ที่กินพืชเป็นอาหาร 2) สัตว์ที่กินสัตว์เป็นอาหาร 3) สัตว์ที่กินทั้งพืชและสัตว์เป็นอาหาร และในสัตว์นั้นยังสามารถนิยามลักษณะเพิ่มเติมได้อีกคือเป็นสัตว์ที่เลี้ยงลูกด้วยนมและสัตว์ที่ออกลูกเป็นไข่ เช่น Natural Language Processing (NLP), Data Mining, Information Retrieval ลักษณะของออนโทโลยี คือโครงสร้างลำดับชั้น (ให้นึกถึง Tree) ที่ใช้ในการจำแนกกลุ่มของวัตถุหรือสิ่งที่เรากำลังสนใจ ตัวอย่างเช่น ออนโทโลยีการเกษตรอาจจะเริ่มจากผลิตผลการเกษตร โดยออนโทโลยีเป็นแนวทางหนึ่งที่ช่วยในการจัดการฐานความรู้และถูกนำมาประยุกต์ใช้กับงานระบบต่างๆไม่ว่าจะเป็นระบบฐานข้อมูลทางด้านชีววิทยา ระบบพาณิชยอิเล็กทรอนิกส์ รวมไปถึงระบบงานต่างๆ ที่พัฒนาขึ้นบนเว็บเชิงความหมายเพื่อช่วยในการจัดเก็บ ค้นคืนความรู้การแลกเปลี่ยนและการนำมาใช่ใหม่

องค์ความรู้ของออนโทโลยี คือ การอธิบายความสัมพันธ์โครงสร้างความรู้ให้อยู่ในรูปแบบลำดับชั้นเชิงวัตถุ เมื่อมีการเปลี่ยนแปลงฐานข้อมูล แล้วองค์ความรู้จะไม่มีการเปลี่ยนแปลง การเพิ่มหรือเปลี่ยนแปลงฐานข้อมูลจะมีผลต่อส่วนของการเชื่อมต่อกับฐานข้อมูล เช่นเดียวกับโครงสร้างข้อมูล องค์ความรู้มีการเปลี่ยนแปลง เมื่อมีการเปลี่ยนแปลงโครงสร้างความรู้ ตัวอย่างงานที่ใช้องค์ความรู้ เช่น BACIIS ในระบบรวบรวมข้อมูล (Data Integration System) องค์ความรู้จะแสดงกลุ่มของคอนเซพที่อยู่ในโดเมน ทำให้การร้องขอข้อมูลถูกกำหนดด้วยคำที่เป็นมาตรฐาน สามารถอ้างถึงได้ในทุกๆ แหล่งข้อมูลที่เกี่ยวข้องกัน การทำงานขององค์ความรู้ในระบบรวบรวมข้อมูล ดังนี้

- ลดความซ้ำซ้อนในการเข้าถึงและการดึงข้อมูลจากแหล่งข้อมูลต่างๆ
- อธิบายข้อมูลในแต่ละแหล่งข้อมูลตามรูปแบบของคำที่กำหนดในองค์ความรู้
- กำหนดโครงสร้างของข้อมูลในแหล่งข้อมูลตามคอนเซพขององค์ความรู้
- สร้างคอนเซพใหม่ๆ ได้โดยมีพื้นฐานจากคอนเซพที่ได้กำหนดไว้แล้ว
- จัดเตรียมกลไกที่สามารถเลือกคอนเซพในแต่ละองค์ความรู้ที่ต้องการร้องขอได้
- แสดงผลลัพธ์ของการร้องขอในรูปแบบมีโครงสร้าง

องค์ประกอบของออนโทโลยี (Components of ontology)

ออนโทโลยีประกอบไปด้วยคอนเซพซึ่งเป็นพื้นฐานสำคัญในการสร้างฐานความรู้ โดยคอนเซพเหล่านี้จัดเรียงอยู่ในลำดับชั้นการถ่ายทอดความสัมพันธ์ และมีคุณสมบัติเฉพาะในแต่ละคอนเซพ โดยสรุปแล้วองค์ประกอบของออนโทโลยีประกอบด้วยส่วนต่าง ๆ ดังต่อไปนี้

• แนวคิด (Concept) : ความคิดทั่วไปหรือนามธรรมในโดเมนที่เราสนใจ เช่น Computer, PC, Notebook, Keyboard เป็นต้น

• คุณลักษณะ (Property) : คุณสมบัติของคอนเซพ เช่น สี น้ำหนัก เป็นต้น

• ความสัมพันธ์ (Relationship) : ความสัมพันธ์เชิงความหมายระหว่างคอนเซพออนโทโลยี โดยส่วนใหญ่จะประกอบไปด้วยความสัมพันธ์ ได้แก่

- is-a คือ ความสัมพันธ์ที่มีคุณสมบัติการถ่ายทอด คุณสมบัติของคอนเซพแม่
- ไปยังคอนเซพลูก เช่น PC is-a Computer ซึ่งอธิบายได้ว่า PC มีคุณสมบัติเป็น Computer
- part-of คือ ความสัมพันธ์ที่หมายถึงการเป็นส่วนประกอบ เช่น Keyboard
- part-of Computer ซึ่งอธิบายได้ว่า Computer จะต้องประกอบไปด้วย Keyboard
- syn-of คือ ความสัมพันธ์ที่แสดงถึงคอนเซพที่มีความเหมือนเชิงความหมายต่อกัน เช่น Data Processor syn-of Computer ซึ่งอธิบายได้ว่า Computer มีความหมายเดียวกันกับ Data Processor สามารถใช้แทนกันได้
- instance-of คือ ความสัมพันธ์ที่แสดงถึงการเป็นตัวแทน หรือสมาชิกของ คอนเซพ เช่น HP instance-of Computer ซึ่งอธิบายได้ว่า HP เป็นคอมพิวเตอร์ประเภทหนึ่ง

นอกจากนี้ออนโทโลยียังประกอบไปด้วยความสัมพันธ์เชิงความหมายอื่น ๆ ที่สอดคล้องกับโดเมน ซึ่งกำหนดโดยผู้เชี่ยวชาญ

• ข้อความอธิบายคอนเซพ (Axiom) เงื่อนไขหรือ ข้อกำหนดเฉพาะในแต่ละคอนเซพเป็นกลไกสำคัญสำหรับใช้ในการอนุมานความรู้ เพื่อการสร้างความรู้ใหม่จากออนโทโลยี ตัวอย่างเช่น คอมพิวเตอร์จะต้องประกอบด้วยซีพียู เป็นต้น

ประเภทของออนโทโลยี

ออนโทโลยีสามารถจำแนกประเภทได้ตามขอบเขตของการประยุกต์ใช้งานได้เป็น 3 ประเภทด้วยกัน คือ

1. ออนโทโลยีแบบสาธารณะ (Generic) เป็นออนโทโลยีที่เป็นจุดรวมของโครงสร้างของความรู้ในเชิงกว้าง มีความเป็นนามธรรมค่อนข้างสูง (Abstract) เป็นความรู้ในเชิงมหภาพ สามารถถูกเรียกใช้งาน (Reuse) คุณลักษณะ ต่าง ๆ ของออนโทโลยีนี้ไปใช้งานได้โดยไม่ต้องสร้างใหม่ ออนโทโลยีสาธารณะมีชื่อเรียกได้หลายลักษณะ เช่น Upper Ontology, Core Ontology หรือ Reference Ontology ตัวอย่างของ Generic Ontology ได้แก่ Cyc-Upper ซึ่งอธิบายถึงสิ่งต่าง ๆ แยกเป็นประเภท ๆ แต่รายละเอียดย่อย (Specialization) ไม่ได้บรรยายไว้

2. ออนโทโลยีสำหรับระบบงานหนึ่ง ๆ (Domain Oriented Ontology) เป็นออนโทโลยีจัดทำขึ้นโดยมีรายละเอียดครอบคลุมเฉพาะในระบบงานหนึ่งๆ ซึ่งจะรวบรวมคุณลักษณะทั่วไปที่สามารถใช้ร่วมกัน (Generalization)ระหว่างกิจกรรมต่าง ๆ (Task) ภายใต้ระบบงานนั้นๆ ตัวอย่างของออนโทโลยีประเภทนี้ ได้แก่ UML (Medicine), TOVE (U.Toronto CANADA), KSL (Ontolingua)

3. ออนโทโลยีสำหรับกิจกรรม (Task Oriented Ontology) เป็นออนโทโลยีที่พัฒนาขึ้นเพื่อตอบสนองการทำงานของกิจกรรมย่อย ๆ (Task)โดยอาศัยคุณลักษณะ generic จาก Ontology Domain Oriented มาช่วยบรรยายลักษณะเชิงสาธารณะ ตัวอย่างของออนโทโลยีสำหรับกิจกรรม ได้แก่ E.coli, Cereal, Mechanical Machine

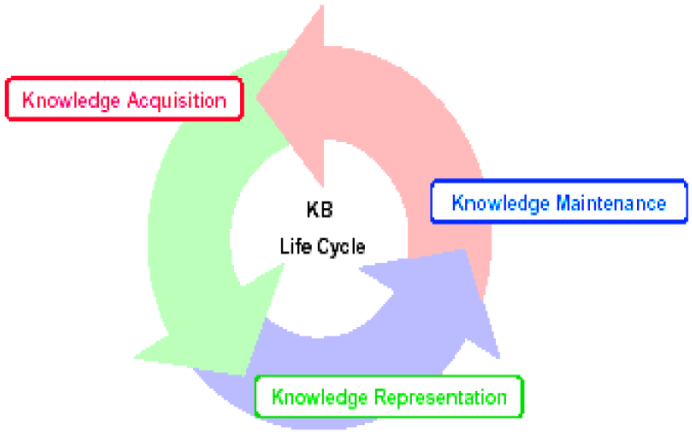
Ontology Library Systems (OLS)

Ontology Library Systems (OLS) เป็นระบบไลบรารีที่มีฟังก์ชั่นต่าง ๆ ที่ช่วยในการจัดกลุ่ม รวบรวม ดูแลรักษา นำส่งและออกเวอร์ชัน ซึ่งออนโทโลยีควรจะง่ายต่อการใช้งาน มีโครงสร้างสำหรับการจัดเก็บและดูแลรักษาออนโทโลยีสามารถปรับเปลี่ยนในการแก้ไข ค้นหา ให้เหตุผล สนับสนุนการนำออนโทโลยีที่มีอยู่กลับมาใช่ใหม่และมีมาตรฐานบนพื้นฐาน Upper Level Ontology และ Ontology Representation Language

การพัฒนาออนโทโลยี (Ontology building)

ออนโทโลยีที่พัฒนาขึ้นจะต้องมีการนิยาม จัดหมวด จัดหมู่ความรู้ (Terminology) ต้องสามารถถูกใช้ซ้ำได้ (Reuse) มีวิธีการจัดการ (Methodology) และสามารถแลกเปลี่ยนกับภายนอก (Interchange) ซึ่งคุณลักษณะต่างๆจะกล่าวในหัวข้อถัดไป ส่วนการจัดการการพัฒนาออนโทโลยีโดยสรุปจะมีขั้นตอนและกิจกรรมที่ต้องกระทำดังนี้

1. วงจรการจัดการความรู้ จะเริ่มต้นจากการทำการรวบรวมจัดสร้างฐานความรู้ (Knowledge acquisition) การใช้งานฐานความรู้ (Knowledge Representation) และการบำรุงรักษา (Knowledge Maintenance)



ภาพประกอบที่ 2-3 ภาพแสดง KB life cycle

2. การจัดการการพัฒนาออนโทโลยี ได้แสดงกระบวนการพัฒนาออนโทโลยีไว้ ประกอบเป็นสามส่วนหลักคือ การบริหารจัดการโครงการ วงจรการพัฒนาออนโทโลยีและการใช้งานบำรุงรักษาออนโทโลยี

3. เฉพาะในส่วนของวงจรการพัฒนาจัดสร้างออนโทโลยี (Ontology life cycle) มีกิจกรรมย่อยๆดังนี้

3.1 การระบุขอบเขต วัตถุประสงค์ (Identify purpose and scope)

3.2 การจัดสร้างออนโทโลยี

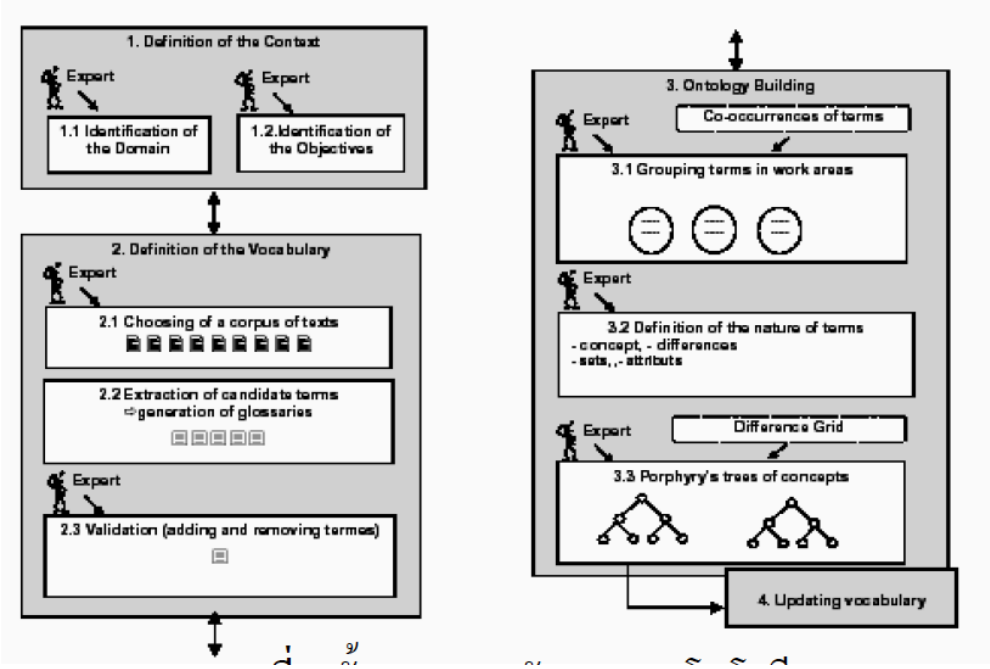
3.2.1 การรวบรวมข้อมูล (Ontology capture)

3.2.2 การพัฒนาโปรแกรม (Ontology coding)

3.2.3 การเชื่อมโยงกับออนโทโลยีอื่นๆ (Integration Existing Ontology)

3.3 การประเมินผล (Evaluation)

3.4 การจัดทำเอกสารคู่มือ (Documentation)



ภาพประกอบที่ 2-4 ขั้นตอนการพัฒนาออนโทโลยี

4. การรวบรวมข้อมูลในส่วนของการรวบรวมข้อมูลจะคล้ายๆกับขั้นตอนการทำ Fact finding techniques ทั่วๆไป แต่เป้าหมายของข้อมูลจะไม่ใช่กระบวนการสถิติปัญหาของระบบงานเหมือนในการวิเคราะห์และออกแบบระบบงานคอมพิวเตอร์ แต่สิ่งที่จะไปรวบรวมจะเป็นคำถาม คำตอบ ต่างๆ (ปัญหา) ที่เกี่ยวข้องกับขอบเขตงานที่กำลังศึกษา

4.1 การรวบรวมข้อมูลจะกระทำด้วยการระดมความคิดจากกลุ่มผู้เชี่ยวชาญ และบุคคลที่เกี่ยวข้อง เพื่อหาคำตอบต่างๆ (Term) ที่เกี่ยวข้อง คำเหล่านี้จะถูกนำมาใช้แทนหน่วยความรู้ (Concept)

4.2 หน่วยความรู้ที่ได้ (Concept or Term) จะนำมาทำการพิจารณาจัดหมวดหมู่ โดยพิจารณาคุณลักษณะร่วม (Common generalization) และคุณลักษณะพิเศษ (Specialization) และพิจารณาความสัมพันธ์เกี่ยวข้องกันระหว่างกัน (Relationship) เทคนิคที่ใช้อาจใช้วิธีการ Topdown approach, Bottom-up หรือ Middle Out approach

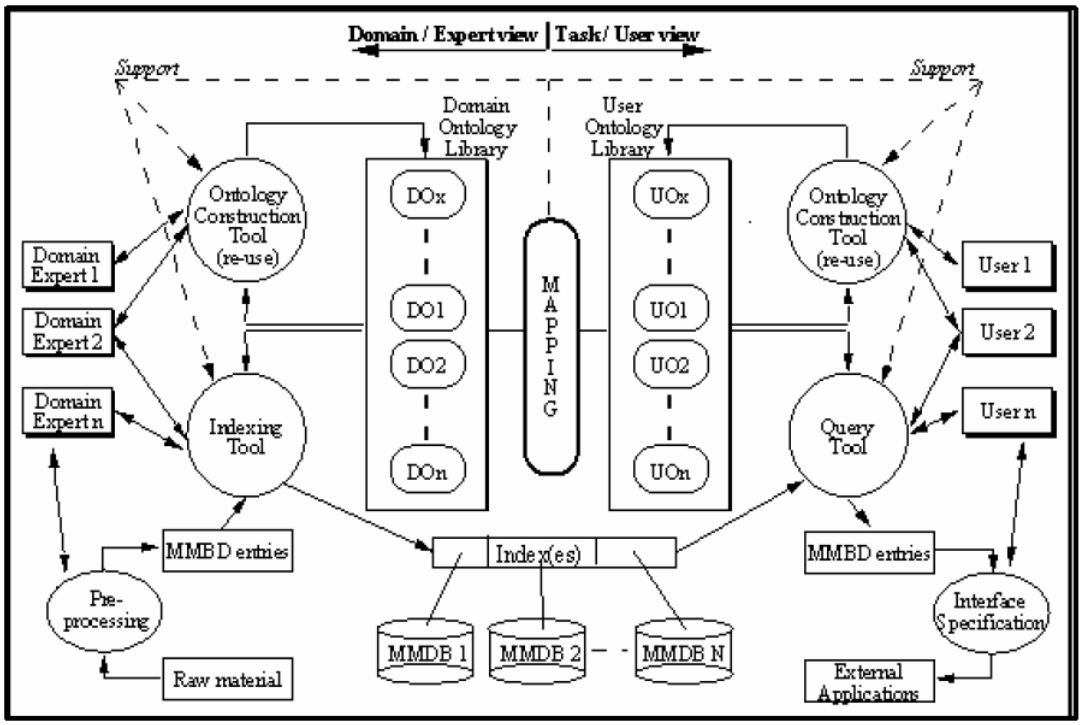
4.3 เมื่อได้ภาพ (Taxonomy) ของ Concepts ต่างๆแล้ว ผู้เชี่ยวชาญจะพิจารณาความครบถ้วน สมบูรณ์อีกหน (ภาพโครงสร้างของตัว Taxonomy จะได้กล่าวในหัวข้อต่อไป)

5. ขั้นตอนการเชื่อมโยงออนโทโลยีอื่นๆ เป็นขั้นตอนที่ผู้พัฒนาออนโทโลยีพยายามรวบรวมออนโทโลยีอื่นๆที่ได้พัฒนาไว้แล้ว โดยเฉพาะ Generic Ontology เพื่อช่วยให้การพัฒนา Domain หรือ Task Ontology เป็นด้วยความรวดเร็ว โดยต้องพิจารณาถึงความคล้ายคลึงกัน (Similarity) โดยเฉพาะอย่างยิ่งด้านความหมาย (Semantic) ในทางกลับกันออนโทโลยีที่พัฒนาขึ้นอาจมีขนาดใหญ่ เพื่อความสะดวกในการจัดการ การอนุญาตสิทธิการใช้งาน (Rights-Trustee) การกระจาย (Refactoring) จะทำให้ได้ออนโทโลยีแยกเป็นกลุ่มก้อนที่เกี่ยวข้องกันสะดวกต่อการจัดการการใช้งานต่อไป

6. ขั้นตอนการโปรแกรมออนโทโลยี (Coding) เป็นขั้นตอนการแปลงการออกแบบโครงสร้าง Concept (Formal conceptualization Taxonomies) เป็นโปรแกรมโดยใช้ภาษาที่พัฒนาขึ้นมาสำหรับการจัดสร้างออนโทโลยี แนวทางในการโปรแกรมจะแตกต่างกันไปขึ้นกับธรรมชาติของการทำ Ontology Formal Conceptualization ซึ่งแยกได้เป็นหลายแนวทาง แนวทางหลักๆที่นิยมใช้กัน ได้แก่ Vocabulary, Frame base (Object model), Logic ซึ่งรายละเอียดจะได้กล่าวในหัวข้อต่อไป

7. ขั้นตอนการประเมิน (Evaluation) เป็นขั้นตอนในการตรวจสอบออนโทโลยีโดยการทดสอบหาคำตอบ จากการป้อนข้อคำถามเข้าไป แล้วพิจารณาความสมบูรณ์ (Completeness), ความตรง (Consistency), ความซ้ำซ้อน (Redundancy), ความเป็นกลุ่ม-หมวดหมู่ที่ดี (Granularity) เป็นต้น

8. สถาปัตยกรรมของฐานความรู้หลังจากการจัดสร้างออนโทโลยี จะทำการติดตั้งใช้งานระบบออนโทโลยีเพื่อนำไปใช้งาน ผู้พัฒนาจะมีการออกแบบสถาปัตยกรรมเพื่อแสดงผังโครงสร้างของระบบงานโดยรวมของฐานความรู้ ซึ่งไม่มีรูปแบบที่ตายตัว ตัวอย่างเช่น



ภาพประกอบที่ 2-5 แสดงตัวอย่างของสถาปัตยกรรมฐานความรู้

จากภาพจะแยกเป็นสองส่วนคือส่วนการพัฒนาบำรุงรักษาออนโทโลยี (Domain/Expert views) ซึ่งทำหน้าที่ในการพัฒนา บำรุงรักษาฐานความรู้เป็นหลัก ส่วนที่สองเป็นส่วนที่เกี่ยวข้องกับผู้ใช้งาน (Task/User views) จะทำการร้องขอฐานความรู้ที่ต้องการผ่านระบบประสานผู้ใช้ (Application interface -ปัจจุบัน IEEE กำลังพัฒนามาตรฐานในส่วนนี้อยู่) ในระบบจะต้องมีส่วนที่ทำหน้าที่ในการเปลี่ยนแปลงคำร้องขอของผู้ใช้ (Query) ซึ่งมักจะเป็นภาษาธรรมชาติ (Natural language) ให้เป็นรูปแบบเชิงตรรก (First order logic:FOL) เพื่อให้สามารถสืบค้น อนุมาน (Inference) ออกเป็นดัชนี (Index) ที่ตรงกับคำตอบในเชิงตรรกะ(Axiom-FOL หรือโครงสร้างของออนโทโลยีที่มี Occurrence รองรับ Lattice) ดัชนีนี้เชื่อมโยงสัมพันธ์กับ Lattice หากหาพบ ค่าดัชนี

จะเป็นตัวอ้างอิง Occurrence ในฐานข้อมูลเชิงสัมพันธ์(RDB)เพื่อดึงระเบียนหรือ Record (Occurrence) ออกมาจัดส่งให้ผู้ร้องขอต่อไป หาก Occurrence ที่จัดเก็บในฐานข้อมูลมีการจัดเก็บในรูปของ Descriptive logic, FOL, KIF การแสดงเป็นประโยคคำตอบอาจไม่สละสลวยจึงต้องมีการพัฒนาตัวกรอง จัดเรียงเพิ่มเติม ถ้อยคำ เพื่อให้ได้ประโยคที่อ่านแล้วเข้าใจได้ดีโดยมนุษย์

ประเภทออนโทโลยีเชิงฟอร์มอล

หลังจากที่ได้ทำการรวบรวม Term หรือ Concept ต่างๆแล้วผู้ออกแบบต้องเลือกทำการระบุการแสดงรูปแบบของ Concept ในลักษณะที่เป็นฟอร์มอล รูปแบบหรือประเภทออนโทโลยี จะมีอยู่หลายรูปแบบ แต่รูปแบบหลัก ๆ ที่เป็นที่นิยมและมีการพัฒนาภาษา Language และ Tool พิเศษมาช่วยงานจะแยกเป็น 3 ประเภทคือ

1. เชิงคำศัพท์ (Vocabularies-Taxonomy) เป็นออนโทโลยีที่ใช้งานได้ดีด้วยวิธีการปฏิบัติ (Hand craft) โครงสร้างของฐานความรู้จะเป็นลักษณะ Hierarchical tree และมีการถ่ายทอดคุณลักษณะ ภายใต้กิ่งย่อย ๆ ลงไปการจัดการฐานความรู้จะสามารถเรียกค้นได้ เฉพาะ Concept ที่ปรากฏอยู่บนกิ่งโครงสร้างเท่านั้น วิธีนี้จึงนิยมใช้กับฐานความรู้ที่มี Term หรือ Concept ที่แน่นอนไม่มีการใช้คำอื่น (Synnonym) และขอบเขตความครอบคลุมความรู้ไม่สลับซับซ้อนมากนัก ความรู้มีลักษณะหยุดนิ่ง (Static) ไม่เปลี่ยนแปลงบ่อยนัก ตัวอย่าง ของออนโทโลยีประเภทนี้ได้แก่ Gene Ontology : Go, Mechanical Machining, CEREAL, etc.

2. เชิงกรอบความรู้ (Frame-based system) ออนโทโลยีมีลักษณะโครงสร้างที่สลับซับซ้อนกว่า โดยแต่ละ Concept จะเรียกเป็น Frames หรือ class แต่ละ Frames จะมี Slot หรือตัวแปร Attribute ระบุ Property ของแต่ละ Concept นอกจากนั้นยังมีส่วนของเงื่อนไขข้อจำกัดของ Property เรียกเป็น Facets นั่นคือหนึ่ง Concept จะมีการอธิบาย Property และ Facets รูปแบบการเขียนจะเขียนในรูปแบบคล้าย ๆ Class diagram ใน UML (Object oriented diagram) การเชื่อมโยงค้นระหว่าง Class จะมี Relation ระหว่างกันซึ่งจะมีใน 2 ลักษณะนี้จะคล้าย ๆ กับ ‘kind-of’, ‘part of’ ใน Object oriented class diagram Relation ประเภทนี้เรียกเป็น Taxonomic relation ส่วนความพันธ์อีกประเภทคือ Non-Taxonomic relation จะเป็น Relation ที่ต่าง class มีต่อกันในลักษณะอื่น ๆ เช่น ‘has a’ เป็นต้น ออนโทโลยีรูปแบบเชิงวัตถุนี้เป็นที่นิยมใช้ในการพัฒนา เนื่องจากคล้ายคลึงวิธีการเชิงวัตถุ ตัวอย่างของออนโทโลยี ที่พัฒนาโดยใช้หลักการนี้ ได้แก่ Ecocyc, Ribo web ฯลฯ ภาษาออนโทโลยีประเภทนี้ได้แก่ Ontolinqua ปัญหาของออนโทโลยีประเภทนี้คือ เงื่อนไข (Constraint) ใน Facets เป็นเงื่อนไขที่ช่วยในการเชื่อมต่อเฉพาะ Class ที่มีลักษณะเป็น Primitive หรือ class ลูก และมีเงื่อนไขที่ค่อนข้างจำกัดการจัดสร้างยังคงต้องใช้ผู้เชี่ยวชาญเข้าพิจารณา (Hand-crafted) เป็นหลัก ข้อจำกัดที่สำคัญที่สุดอีกข้อ คือ การอนุมาน (Inference) จะกระทำได้น้อยลง ตาม

ข้อจำกัดของเงื่อนไขจึงได้มีผู้พัฒนาแนวคิด Object constraint language : OCL เพื่อช่วยเสริมเงื่อนไขให้มีความยืดหยุ่นขึ้น การอนุมานความรู้ใหม่ ๆ ที่ต่างจากที่กำหนดไว้ สามารถกระทำได้ง่ายขึ้น แต่ก็ยังไม่สมบูรณ์พอ

3. ออนโทโลยีเชิงตรรกพรรณนา (Descriptive Logics) เป็นออนโทโลยีที่มีการนำเอาความรู้ ที่มาจากการรวบคำถาม-คำตอบ มาหา Concept และบทบาท (role) ที่ Concepts มีต่อกัน โดยพรรณนาความเกี่ยวเนื่องระหว่างกันออกมา โดยใช้สัญลักษณ์ทางตรรกะ เริ่มต้นจากข้อความเช่น “A cat is on a mat” จะนำข้อความมาแสดงภาพเป็นสัญลักษณ์ Conceptual graphs: CG

ภาพประกอบที่ 2-6 ภาพแสดง Conceptual graph

จากนั้นทำการแปลงรูปให้เป็น Linear form (LF) [CAT] -> (On)-> [MAT] แล้วทำการแปลง LF เป็นรูปแบบทางตรรกะ Conceptual Graph Interchange Format : CGIF ดังนี้ (exists((? x cat) (? Y Mat)) (On ? x ? y)) รูปแบบของ CGIF จะเป็นกึ่ง ๆ มาตรฐานกลาง ซึ่งจะเหมือนกับพิมพ์เขียว เพื่อนำไปพัฒนาเป็นออนโทโลยีโดยภาษาออนโทโลยี (ที่มีรูปแบบ Descriptive logics) ต่อไป ตัวอย่างของ CGIF ที่ได้จัดทำขึ้นคือ KIF (Knowledge Interchange Format) ส่วนมากแล้วภาษาต่าง ๆ จะแปล CGIF หรือ KIF ให้เป็น Predicate Calculus ซึ่งเป็นตรรกะ ระดับล่าง (Primitive) สุดที่ใช้ในการแสดงตรรกที่เป็นรูปแบบทางการ (Formal) จากตัวอย่าง ข้างต้น Predicate calculus จะเท่ากับ

$$(\exists x{:}\mathrm{Cat})(\exists y{:}\mathrm{Mat})\mathrm{on}(x,y)$$

ภาพประกอบที่ 2-7 Predicate calculus

หากไม่ใช้ภาษา Ontology ผู้พัฒนาอาจใช้ภาษาโปรแกรมลอจิกที่รู้จักกันดีได้แก่ภาษา Prolog KIF หรือPredicate Calculus นี้ถือเป็น Formal Axiom ของ Concept บน Relation/Role ที่มีต่อ Concept หรือแม้แต่กับ Role อื่น ๆ ในเชิงตรรกะ หากมีคำถามที่แตกต่าง จากต้นกำเนิด (A cat is on a mat) หน่วยประมวลผลจะทำการพิจารณาหา Predicate calculus ของ Query sentences ใหม่ แล้วนำไปอนุมานกับ Pre-predicate calculus เพื่อหาความรู้ในเส้นทางที่แตกต่างกันไปได้ (Lattice) นั่นคือออนโทโลยีแบบตรรกพรรณา สามารถอนุมาน (Inference, Deduce, Render) ความรู้ใหม่ ๆ ที่แตกต่างจากที่ได้กำหนดไว้เดิมได้คล้ายความสามารถของมนุษย์ที่สามารถอนุมานประเภทของสิ่งของเพื่อศึกษา แยกแยะ แจกแจง จัดหมวดหมู่คาดเดาคุณลักษณะโดยการอนุมานพฤติกรรม (role)และคุณลักษณะของแต่ละ concept ตัวอย่างของออนโทโลยี แบบเชิงตรรกพรรณนา ได้แก่ TOVE

การเชื่อมโยงออนโทโลยี (Ontology Integration)

ออนโทโลยี หรือ ฐานความรู้ของกิจกรรม (Task) ต่าง ๆ ในองค์กรอาจมีการพิจารณานำมารวมกัน เป็น ฐานความรู้ขององค์กร หรือแม้แต่การพิจารณา เชื่อมโยงรวมออมโทโลยีที่พัฒนาขึ้น เพื่อให้เกิดความสมบูรณ์ (Perfect) หรือขยายของขอบเขต (Coverage) ได้มากขึ้น เป็นสิ่งที่นักออกแบบฐานความรู้ความสนใจ ลักษณะการกระทำจะคล้าย ๆ กับการทำการรวมเชื่อมโยงของฐานข้อมูลสกีมาหลาย ๆ ฐานข้อมูลเข้าด้วยกันและจะมีความยุ่งยากกว่า เพราะ Concept ต่างๆ มีเงื่อนไข (Facets, Axiom) ต่างกัน สิ่งที่จะรวมกันได้ต้องเจาะจงที่เรื่องเดียวกัน ในทิศทางที่สอดคล้องกัน (Consistency) แต่ถ้าเกิดความขัดแย้งขึ้น (Semantic conflict) จะต้องมีการพิจารณาเลือกจัดการขจัดความขัดแย้ง นอกจากความขัดแย้งในเชิงความเข้าใจความหมายแล้วยังมีปัญหาอื่นๆ อีก ได้แก่

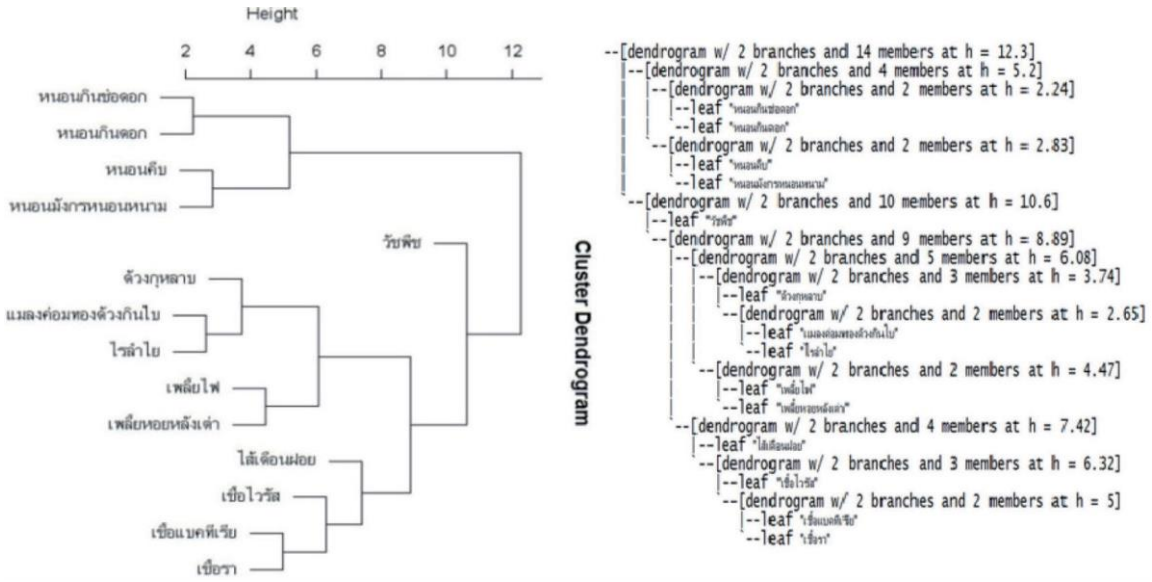
- การใช้แนวทางในการพัฒนาที่แตกต่างกัน เช่น Vocabulary, Frame based, Descriptive logic ซึ่งมีความแตกต่างกันมาตั้งแต่ขบวนการทำ Knowlegde Acquisition ความแตกต่างกันของ Application domain วิธีการปฏิบัติแนวทาง เป้าหมาย แตกต่างกัน
- ความแตกต่างกันของ Application domain การพิจารณาการแก้ไขเพื่อให้สามารถเชื่อมโยงกันได้ต้องใช้ผู้เชี่ยวชาญทำการพิจารณาปรับแก้ส่วนปัญหาด้านแนวทางในการพัฒนาออนโทโลยีนั้นได้โดย กระทำผ่านมาตรฐานการแลกเปลี่ยนร่วมกัน (KIF,CGIF)

การบำรุงรักษาออนโทโลยี(Ontology Maintenance)

ออนโทโลยีที่พัฒนาขึ้นจะเป็นข้อมูลฐานความรู้ที่รวบรวมจากช่วงเวลาหนึ่ง ๆ หากสถานการณ์สิ่งแวดล้อมเปลี่ยนแปลงไปฐานความรู้อาจต้องถูกปรับปรุง (Update) ถูกเพิ่มเติมความรู้ใหม่ ๆ ปัญหาที่จะพบคือ การปรับปรุงแก้ไข Concept, Relation จะกระทบต่อส่วนอื่น ๆ หรือไม่ และหากจะเพิ่มความรู้ใหม่ๆ เข้าไปจะไปทำการเชื่อมต่อเข้าตรวจจุดใดบ้างทั้งนี้เพื่อไม่ให้ฐานความรู้เดิมเปลี่ยนแปลงไป และสามารถสืบค้นความรู้ใหม่ ๆ ได้อย่างมีประสิทธิภาพการบำรุงรักษาออนโทโลยี ใช้เทคนิคการวิเคราะห์ข้อมูลต่าง ๆ เช่น Clustering, Hidden Markov Model, Natural language Analysis, Lexical Conceptual graph โครงการที่ได้ทำการวิจัยต่างพยายามทดลองหาแนวทางที่เหมาะสมใช้ในการบำรุงรักษาออนโทโลยี สรุปแนวทางการบำรุงรักษาออนโทโลยีจะได้ผลดี หากออนโทโลยีเดิมมีการพัฒนาบนมาตรฐานการแลกเปลี่ยนข้อมูลร่วมกัน และจะมีความถูกต้องสูงเป็นออนโทโลยีที่มีขนาดไม่ใหญ่มากนัก หรือเป็นออนโทโลยีทีผ่านกระบวนการขัดเกลาให้มีความซับซ้อนน้อยลงโดยวิธี Canonical graph หรือวิธี Formal Concept Analysis ซึ่งจะกระทำในช่วงการพัฒนาออนโนโลยี หากออนโทโลยีได้รับการจัดทำโครงสร้างแล้ว (Taxonomy) ปรากฏว่าออนโทโลยีมีโครงสร้างใหญ่ควรทำการแตกเป็นออนโทโลยีย่อย (Refactoring) ให้เป็นออนโทโลยีขนาดเล็กลง คล้าย ๆ กับการแตกโปรแกรมขนาดใหญ่ให้เป็น Module หรือ Procedure ย่อย ๆ เป็นต้น

การจัดกลุ่มแบบลำดับชั้น (Hierarchical Clustering Algorithm)

การจัดกลุ่มแบบลำดับชั้นเป็นหนึ่งในหลายๆวิธีที่ ใช้สำหรับจัดกลุ่มของข้อมูล โดยจะทำการจัดการข้อมูลให้อยู่ในรูปของลำดับชั้น (Hierarchy) ผลของการจัดกลุ่มด้วยวิธีการนี้มักจะให้ผลลัพธ์ออกมาเป็นเดนโดรแกรม (Dendrogram) โดยที่โหนดรากจะแสดงถึงสิ่งที่ใช้กล่าวอ้างถึงกลุ่มของข้อมูลนั้นหมดและโหนดจะถือเป็นตัวข้อมูล โดยทั่วไปถูกแบ่งออกเป็นสองประเภท ได้แก่ Agglomerative method และ Divisive method โดยการทดลองการจัดกลุ่มนี้ได้เลือกเนื้อหาที่เกี่ยวข้องกับโดเมน X มาเพื่อมาจัดกลุ่มโดยมีเนื้อหาที่เกี่ยวข้องกับข้อมูลที่เกี่ยวข้องกับ X ซึ่งในชั้นต้นเอกสารที่นำมาทดสอบจะต้องผ่านการประมวลผลการตัดคำโดยใช้ SWATH (Smart Word Analysis for Thai) หลังจากนั้นจะทำการคัดเลือกคำจากคีย์เวิร์ดที่เกี่ยวข้องกับโดเมนที่กำหนด จากนั้นจะทำการวิเคราะห์หาความถี่ของคำที่เกิดขึ้นในเอกสาร แล้วรวบรวมให้อยู่ในตาราง tf-idf เพื่อนำมาจัดกลุ่มด้วยเทคนิคแบบลำดับ ชั้นด้วยภาษาอาร์ (R language) โดยทำการหาระยะห่างก่อน



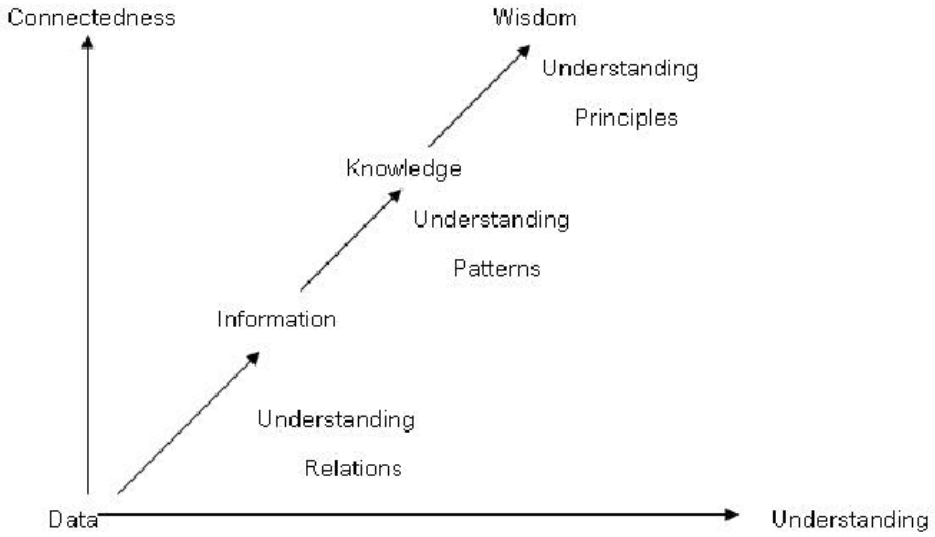
ภาพประกอบที่ 2-8 ตัวอย่างเดนโดรแกรมจากการจัดกลุ่มแบบลำดับชั้น

การแปลงโครงสร้างให้อยู่ในรูปของภาษาเว็บออนโทโลยี

หลังจากที่ได้ผลลัพธ์จากการจัดกลุ่มด้วยเทคนิคแบบลำดับชั้นมาใส่ในโครงสร้างภาษาเว็บออนโทโลยี โดยใช้ภาษาเพิร์ล (Perl Language) ช่วยในการเตรี ยมโครงสร้างภาษาเว็บออนโทโลยี จากนั้นนำภาษาเว็บ ออนโทโลยีที่ได้ไปสร้างเป็นออนโทโลยีด้วยโปรแกรม โปรเตเจ้ (Protege) ซึ่งเป็นเครื่องมือที่ใช้ในการสร้าง ออนโทโลยีชนิดหนึ่ง (Ontology Editor) โดยทำการ นำเข้าไฟล์ภาษาเว็บออนโทโลยีเข้ามาในโปรแกรมเพื่อให้โปรแกรมทำการอ่านโครงสร้างภาษา จากนั้นโปรแกรมจะทำการแสดงออนโทโลยี จากกระบวนการสร้าง ออนโทโลยีแบบกึ่งอัตโนมัติ ในงานวิจัยนี้ทำให้สามารถ สกัด รวบรวมและจัดกลุ่มความรู้ โดยอาศัยการจัด กลุ่มแบบลำดับชั้นเข้ามาช่วยในการจัดกลุ่มโครงสร้าง ข้อมูลให้อยู่ในรูปแบบโครงสร้างที่เป็นลำดับชั้นตาม แนวคิดของออนโทโลยีได้อย่างรวดเร็ว ซึ่งจะเป็นการ เพิ่มประสิทธิภาพในการสร้างออนโทโลยีสำหรับการนำเสนอข้อมูลองค์ความรู้ต่อไป

การจัดการความรู้เชิงความหมาย (Semantic-based Knowledge Management)

ความรู้ (knowledge) คือ ผลสรุปของการสังเคราะห์สารสนเทศ (information) โดยพิจารณาถึงความสัมพันธ์ของสารสนเทศเทียบเคียงกับความรู้เดิมที่มีอยู่จนได้ผลสรุปที่ชัดเจนถูกต้อง สามารถนำไปประยุกต์ใช้ในกิจกรรมต่างๆ ต่อไปได้อย่างเหมาะสม จากนิยามดังกล่าวอาจสรุปได้ว่าสิ่งสำคัญที่ทำให้ความรู้แตกต่างจากสารสนเทศก็คือความรู้เกิดจากการสั่งสมและสังเคราะห์จากสารสนเทศปริมาณมาก จนเกิดเป็นข้อสรุปที่สามารถนำไปใช้ให้เกิดประโยชน์ได้



ภาพประกอบที่ 2-9 สเปคตรัมของความรู้ (Knowledge spectrum)

รูปแบบการจัดการความรู้ในปัจจุบันอาจแบ่งได้เป็น 2 ประเภทใหญ่ คือ

1. การจัดการความรู้ที่ชัดแจ้ง หรือ ความรู้สาธารณะ (Explicit Knowledge) มักอยู่ในรูปแบบของการจัดเก็บ และสืบค้นข้อมูลจากเอกสาร หรือ ฐานข้อมูลที่มีอยู่

2. การจัดการความรู้ที่ซ่อนเร้น หรือ ความรู้ที่อยู่ในตัวบุคคล (Tacit Knowledge) มักอยู่ในรูปแบบของการส่งเสริมให้เกิดการบันทึก ถ่ายทอดข้อมูลความรูของแต่ละบุคคลให้กับผู้อื่น เช่นในรูปแบบของการเขียนบล็อกหรือวิกิ เป็นต้น

การจัดการความรู้เชิงความหมาย (Semantic Knowledge Management) จึงเป็นรูปแบบการจัดการความรู้ในอีกรูปแบบหนึ่ง ที่มุ่งเน้นการจัดเก็บองค์ความรู้ที่สามารถนำไปใช้งานในโปรแกรมคอมพิวเตอร์ได้ ในรูปแบบของฐานความรู้สำหรับโปรแกรมคอมพิวเตอร์หรือออนโทโลยี (Ontology) ซึ่งวิธีการหนึ่งที่ใช้ในการพัฒนาคือโดยการใช้กระบวนการทางวิศวกรรมความรู้ (Knowledge Engineering)

การจัดการความรู้เชิงความหมายจำเป็นต้องอาศัยแหล่งความรู้ที่มีอยู่ ทั้งที่อยู่ในรูปแบบของเอกสารอ้างอิง (Reference documents) และจากผู้เชี่ยวชาญเฉพาะสาขา (Domain experts) ดังนั้นการจัดการความรู้เชิงความหมายจึงต้องมีการผสมผสานทั้งการจัดการความรู้ที่ชัดแจ้ง (Explicit Knowledge) และการจัดการความรู้ที่อยู่ในตัวบุคคล (Tacit Knowledge) เข้าด้วยกัน

หากพิจารณาตามทิศทางการพัฒนาเทคโนโลยีเว็บแล้ว จะพบว่าเทคโนโลยีเว็บในยุคเริ่มต้นมุ่งเน้นที่การจัดการความรู้ที่ชัดแจ้ง ในรูปแบบของการจัดเก็บและสืบค้นข้อมูลจากเอกสาร HTML และ ฐานข้อมูลจากเว็ปไซต์ต่างๆ เป็นหลัก ในขณะที่เว็บยุคที่ 2 มุ่งเน้นที่การจัดการความรู้ที่อยู่ในตัวบุคคลมากยิ่งขึ้น ในรูปแบบของการเขียนบล็อก และวิกิ รวมทั้งเว็ปไซต์เครือข่ายสังคม ดังเช่น เว็บ Hi5, Facebook และ Twitter เป็นต้น ในเว็บยุคถัดไปจะมุ่งเน้นที่การจัดการความรู้เชิงความหมายมากยิ่งขึ้น เพื่อนำไปสู่การพัฒนาโปรแกรมตัวแทนที่มีความชาญฉลาด (Intelligent Agents) เพื่อมาช่วยในการประกอบกิจกรรมและสืบค้นข้อมูลของผู้ใช้ได้ดียิ่งขึ้น

เทคโนโลยี	การจัดการความรู้	การประยุกต์ใช้งาน
Web 1.0	Explicit Knowledge Management	HTML Documents, Web-Database
Web 2.0	Tacit Knowledge Management	Wiki, Blogs, Social Networks
Web 3.0	Semantic Knowledge Management	Intelligent Agents, Question-Answering

ภาพประกอบที่ 2-10 การจัดการความรู้เชิงความหมาย (Semantic Knowledge Management) สำหรับเทคโนโลยีเว็บยุคหน้า

ฐานความรู้สำหรับโปรแกรมคอมพิวเตอร์หรือออนโทโลยี

ฐานความรู้สำหรับโปรแกรมคอมพิวเตอร์ หรือออนโทโลยี (Ontology) เป็นรูปแบบองค์ความรู้เฉพาะทาง (Domain Knowledge) ที่ส่วนใหญ่จะเกิดจากการพัฒนาขึ้นโดยวิศวกรความรู้ (Knowledge engineers) ร่วมกับผู้เชี่ยวชาญเฉพาะทาง (Domain experts) โดยมีวัตถุประสงค์หลักคือเพื่อให้สามารถนำความรู้เฉพาะทาง ไปประยุกต์ใช้ในโปรแกรมคอมพิวเตอร์ได้หลากหลายชนิด เทคโนโลยีวิศวกรรมความรู้ (Knowledge engineering) มีส่วนสำคัญอย่างยิ่งต่อการพัฒนา จัดเก็บ และแบ่งปันองค์ความรู้จากผู้เชี่ยวชาญเฉพาะทางให้สามารถนำไปใช้งานได้ในโปรแกรมและระบบคอมพิวเตอร์ต่างๆ ให้สามารถทำงานได้อย่างชาญฉลาดและมีความเป็นอัตโนมัติมากยิ่งขึ้น ประโยชน์ของการพัฒนาและประยุกต์ใช้งานฐานความรู้สำหรับโปรแกรมคอมพิวเตอร์ อาจสรุปได้เป็น 4 ประการหลัก คือ

1. เพิ่มความอัตโนมัติของกระบวนการ (Automation)
2. ลดภาระของมนุษย์ (Reduced workloads)
3. เพิ่มความแม่นยำ ลดข้อผิดพลาดในกระบวนการทำงาน (Reduced errors)
4. สามารถนำไปประยุกต์ใช้งานได้ในโปรแกรมและระบบสารสนเทศต่าง ๆ ได้กว้างขวางยิ่งขึ้น (Interoperability) เนื่องจากฐานความรู้สามารถแบ่งปันและใช้ซ้ำได้ (share and reuse)

มาตรฐานของภาษาที่ใช้ในการพัฒนาออนโทโลยี (Ontology language) เพื่อให้สามารถแบ่งปันและแลกเปลี่ยนข้อมูลได้บนเครือข่ายเว็บ คือ มาตรฐาน OWL (Web ontology language) ซึ่งเป็นภาษามาตรฐานสำหรับการพัฒนาออนโทโลยีเพื่อการใช้งานตามแนวทางของเว็บความหมาย โดยภาษาดังกล่าวได้ถูกพัฒนาขึ้นโดยอิงจากมาตรฐาน RDF (Resource description framework) โดยใช้รูปแบบภาษา XML (Extensible markup language) และมาตรฐานการอ้างอิงข้อมูล URIโดยมาตรฐานเหล่านี้ล้วนเป็นองค์ประกอบสำคัญต่อการพัฒนาเว็บความหมาย

เทคโนโลยีที่เกี่ยวข้อง

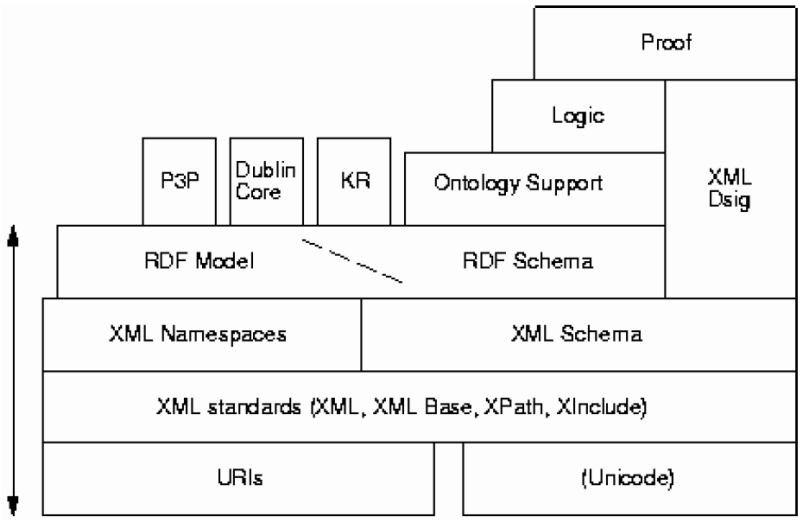
ภาษาที่ใช้ในการพัฒนาออนโทโลยีภาษาที่ใช้ในการพัฒนาออนโทโลยีจะจำแนกตามประเภทของออนโทโลยีเชิงฟอร์มอล

- เชิงกรอบความรู้ ได้แก่ ภาษา Ontolingua, Frame Ontology, DAML - OIL , OWL ฯลฯ
- เชิงพรรณาตรรก ได้แก่ Prolog, KIF, FACT ฯลฯ

นอกจากการพัฒนาออนโทโลยีด้วยภาษาต่าง ๆ ข้างต้น ได้มีการพัฒนาโปรแกรมเอดิเตอร์ (Ontology Editor) เพื่อช่วยในการพัฒนาใช้งานออนโทโลยี

- เชิงกรอบความรู้ ได้แก่ MBO, GKB, Protege
- เชิงพรรณนาตรรก ได้แก่ OIL-ED (Manchester U. UK)

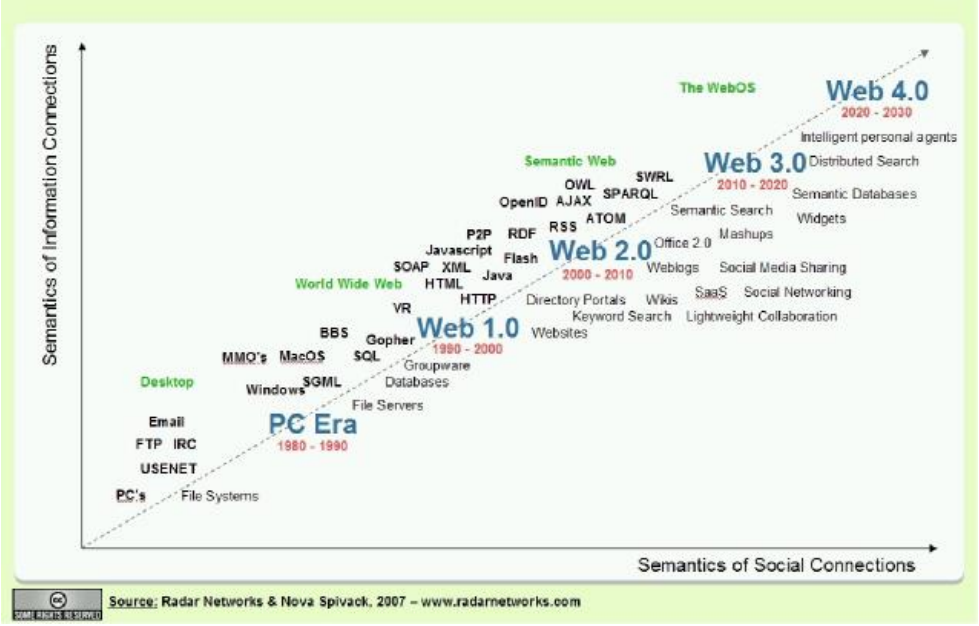
สำหรับการพัฒนาออนโทโลยี ให้ทำงานได้บน Platform ของ Internet จะมีการพัฒนาภาษาและ Tool ช่วยงาน ได้แก่ OWL (Web ontology Language), RDF (Resource definition frame work) สำหรับ RDF เป็นภาษาที่ช่วยในการสร้าง Metadata มาครอบหรือให้ข้อมูลของออนโทโลยี (Metadata) เพื่อให้สามารถสืบค้น Semantic ของ ฐานความรู้ออนโทโลยีได้ แต่หากต้องการการสืบค้น Semantic โดย Metadata ผ่าน RDF ของฐานความรู้ให้สามารถ Interoperation ได้ที่ต่าง Platform ผู้พัฒนาจะต้องทำการใส่Metadata ด้วยภาษา XML เพิ่มอีกหนึ่งภาษาดังภาพ



ภาพประกอบที่ 2-11 ภาพแสดง Metadata ของ Ontology

การสร้างออนโทโลยีนั้นจะต้องอาศัยภาษาเว็บออนโทโลยี (Ontology Web Language: OWL) สำหรับอธิบาย โครงสร้างของความรู้อยู่บนพื้นฐานของภาษาอาร์ดีเอฟ (Resource Description Framework: RDF Language) ซึ่งใช้หลักไวยากรณ์ตามภาษาเอ็กซ์เอ็มเอล (XML extensible Markup Language) ซึ่งเป็นภาษาที่เป็น มาตรฐาน และถูกรับรองโดยองค์กร W3C ซึ่งยังคงใช้ รูปแบบและไวยากรณ์ตามแบบภาษาอาร์ดีเอฟ ถูกสร้าง ขึ้นมาเพื่อรองรับการทำออนโทโลยีโดยเฉพาะ โดยมีวัตถุประสงค์เพื่อใช้เป็นภาษามาตรฐานสำหรับเว็บเชิงความหมาย

เว็บความหมาย (Semantic Web) หากมองย้อนกลับไปถึงจุดเริ่มต้นของคอมพิวเตอร์ส่วนบุคคลเมื่อประมาณ 30 ปีที่แล้ว (ทศวรรษ 1980) จะพบว่าข้อมูลจะถูกจัดเก็บในรูปของไฟล์ในเครื่องคอมพิวเตอร์ของแต่ละบุคคล โดยมิได้มีการแบ่งปันกันกว้างขวางนัก จนเมื่อเข้าสู่ทศวรรษที่ผ่านมา (ทศวรรษ 1990) การเกิดขึ้นของเครือข่ายอินเตอร์เน็ตได้ส่งผลให้เกิดการเปลี่ยนแปลงครั้งใหญ่เข้าสู่ยุคเทคโนโลยีสารสนเทศ หรือ IT ก่อให้เกิดการสร้างและแบ่งปันข้อมูลสารสนเทศกันอย่างกว้างขวาง ในรูปแบบของเอกสาร HTML ที่มีการเชื่อมโยงกันบนเครือข่ายเว็บ จนเมื่อเข้าสู่ทศวรรษนี้ (ทศวรรษ 2000) เราได้เห็นการเกิดขึ้นของรูปแบบการสร้าง และแบ่งปันความรู้ในรูปแบบของวิกิ และ บล็อกเกิดขึ้นอย่างกว้างขวาง ดังเช่น สารานุกรมเสรีวิกิพีเดีย



ภาพประกอบที่ 2-12 ทิศทางการพัฒนาเทคโนโลยีเว็บ

(en.wikipedia.org) หรือ เว็บไซท์ Gotoknow (Gotoknow.org) หรือ เว็ปไซท์ทวิตเตอร์ (Twitter.com) เป็นต้น ก่อให้เกิดการแบ่งปัน แลกเปลี่ยนเรียนรู้ผ่านเครือข่ายเว็บ ซึ่งหลายคนเรียกกันว่าเป็นเทคโนโลยีเว็บยุคที่ 2 หรือ เว็บ 2.0 นั่นเอง

เว็บความหมาย หรือ Semantic Web เป็นวิสัยทัศน์ของทิศทางการพัฒนาเทคโนโลยีเว็บในยุคถัดไป หรือเทคโนโลยีเว็บ 3.0 ซึ่งคาดว่าจะเกิดขึ้นในทศวรรษหน้า (ทศวรรษ 2010) ที่ข้อมูลมีการเชื่อมโยงกันมากยิ่งขึ้น ในลักษณะของเครือข่ายเชิงความหมาย (Semantic Network) เพื่อนำไปสู่การพัฒนาโปรแกรมประยุกต์ที่มีความชาญฉลาดมากยิ่งขึ้น เช่น โปรแกรมตัวแทนอัจฉริยะ (Intelligent Agent) การสืบค้นข้อมูลที่อิงตามความหมาย (Semantic Search) เป็นต้น (รูปที่ 5) โดยมีหน่วยงาน W3C (http://www.w3.org/) เป็นองค์กรสากลที่เป็นผู้กำหนดแนวทางการพัฒนาและมาตรฐานสำหรับข้อมูลบนเว็บความหมาย

โดยแนวความคิดของเว็บเชิงความหมาย (Semantic Web) เป็นแนวความคิดของ Sir Tim Berners-Lee ซึ่งเป็นบุคคลที่คิดค้น WWW, URIs, HTTP และ HTML โดยได้แรงสนับสนุนจากองค์กร W3C (World Wide Web Consortium) โดยมีจุดมุ่งหมายเพื่อช่วยให้สามารถค้นหาข้อมูลบนอินเตอร์เน็ตได้อย่างมีประสิทธิภาพมากขึ้นและยังสามารถสร้างความสัมพันธ์ให้กับข้อมูลที่มีมาจากแหล่งข้อมูลที่ต่างกัน ซึ่งจะก่อให้เกิดเป็นฐานข้อมูลที่เชื่อมโยงกันทั่วโลก

ปัจจุบันนี้ทรัพยากรที่อยู่บน WWW (World Wide Web) นั้นมีจำนวนมากมายมหาศาล ซึ่งก่อให้เกิดปัญหาที่สำคัญคือความแม่นยำในการสืบค้นด้วย Search Engine ลดลง กล่าวคือผลลัพธ์ที่ได้จากการสืบค้นด้วย Search Engine มีจำนวนมาก มีทั้งผลลัพธ์ที่ผู้ใช้ต้องการและผลลัพธ์ที่ผู้ใช้ไม่ต้องการ ส่งผลให้ผู้ใช้ต้องใช้เวลาเลือกเอกสารที่ตนเองต้องการ ที่เป็นเช่นนี้เนื่องจากเอกสาร HTML ไม่สามารถอธิบายความหมายสำหรับข้อมูลที่อยู่ภายใน HTML ได้และไม่สามารถเชื่อมโยงความสัมพันธ์ของข้อมูลได้

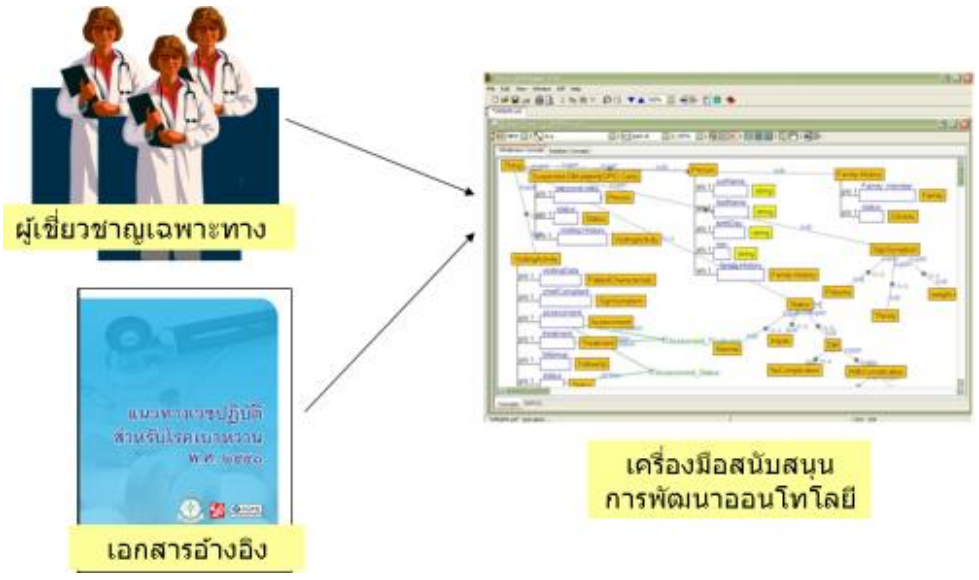
การที่จะทำให้แนวความคิดของเว็บเชิงความหมายเกิดขึ้นได้จริงนั้น จำเป็นต้องมีโครงสร้างของข้อมูลและหลักเกณฑ์ที่ดีสำหรับ Spider หรือ Crawling ที่จะการท่องไปตามเว็บต่างๆ เพื่อเก็บข้อมูลบนเว็บไซต์ สามารถเข้าใจความหมายของข้อมูลและเชื่อมโยงความสัมพันธ์ของข้อมูลได้ ดังนั้น XML และ RDF จึงเป็นเทคโนโลยีที่มีความสำคัญสำหรับแนวความคิดเว็บเชิงความหมาย โดยจะมีการใช้ XML ในการอธิบายโครงสร้างของข้อมูลและใช้ RDF ในการอธิบายความหมายของข้อมูล

อย่างไรก็ตามแนวความคิดเว็บเชิงความหมายยังอยู่ในระหว่างการศึกษาวิจัยและพัฒนา เนื่องจากยังมีความท้าทายหลายประการ อาทิ จะพัฒนาความร่วมมือระหว่างผู้พัฒนาเว็บไซต์ในการเตรียมข้อมูลให้เป็นไปตามมาตรฐานของ Semantic Web อย่างไร และจะจัดการกับปัญหาเกี่ยวกับการละเมิดลิขสิทธิ์ได้อย่างไร เมื่อข้อมูลทุกอย่างทุกเชื่อมโยงเข้าด้วยกัน

เทคโนโลยีเว็บเชิงความหมาย มีองค์ประกอบที่ถูกแบ่งออกเป็นระดับชั้น โดยข้อมูลเชิงความหมาย (Semantic-based Information) คือข้อมูลที่ถูกอธิบายด้วยการกำหนดความหมายไว้เป็น

อย่างดี (Well-Defined Meaning) และข้อมูลเหล่านั้นสามารถอ้างอิงด้วย URI/URL และถูกอธิบายด้วยภาษาเชิงความหมาย (Semantics-Based language) เช่น XML, RDF และ RDFS โดยในการอธิบายข้อมูลเชิงความหมาย สามารถกำหนดคำ (Vocabulary) คุณสมบัติ (Property) และเงื่อนไข (Constraint) ต่างๆ สำหรับการอธิบายข้อมูลซึ่งถูกกำหนดไว้ในออนโทโลยี (Ontology) มาอธิบายในการตีความข้อมูลเชิงความหมายต้องใช้ตรรกะ (Logic) ต่าง ๆ มาพิจารณาในการตีความกระบวนการของการตีความได้ถูกจัดให้อยู่ในชั้นของการพิสูจน์ (Proof) ซึ่งใช้การอนุมาน (Inference) เป็นกลไกของการทำงาน และในชั้น Trust เป็นการสร้างความเชื่อถือได้ในการเผยแพร่และการนำข้อมูลไปใช้งาน อีกทั้งมีการกำหนดการรักษาความปลอดภัยให้กับข้อมูลที่มีการรับส่งระหว่างโปรแกรมประยุกต์ด้วยเทคโนโลยี Digital signature

โปรแกรมเครื่องมือสำหรับสนับสนุนการพัฒนาออนโทโลยี (Ontology Editor) ในปัจจุบันที่ได้รับความนิยม เช่น โปรแกรม Protege ซึ่งพัฒนาโดย มหาวิทยาลัยสแตนฟอร์ด (Stanford University) – http://protege.stanford.edu/, โปรแกรม Hozo ซึ่งพัฒนาโดยมหาวิทยาลัยโอซากา (Osaka University) – http://www.hozo.jp/ เป็นต้น โดยเครื่องมือเหล่านี้เป็นเครื่องมือสนับสนุนกระบวนการวิศวกรรมความรู้ ที่ช่วยให้ผู้ใช้ที่เป็นวิศวกรความรู้ หรือผู้เชี่ยวชาญเฉพาะสาขาสามารถถ่ายทอดและจัดเก็บองค์ความรู้ในรูปแบบของออนโทโลยีได้สะดวก และง่ายดายยิ่งขึ้น



ภาพประกอบที่ 2-13 การพัฒนาฐานความรู้ด้านการดูแลรักษาโรคเบาหวานโดยใช้เครื่องมือสนับสนุนการพัฒนาออนโทโลยี (Ontology Editor)

ตัวอย่างการประยุกต์ใช้งานออนโทโลยี การพัฒนาออนโทโลยีสำหรับการจัดการความรู้ด้านการดูแลรักษาโรคเบาหวาน เป็นความร่วมมือระหว่างหน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา เนคเทค และสถาบันวิจัยและประเมินเทคโนโลยีทางการแพทย์ กรมการแพทย์ เพื่อนำองค์ความรู้ด้านการดูแลรักษาผู้ป่วยโรคเบาหวานที่เป็นระบบ มีหลักฐานทางวิชาการรองรับ (Evidence-based practice guideline) ไปประยุกต์ใช้งานในรูปแบบของระบบแจ้งเตือนความจำทางคลินิกสำหรับฐานข้อมูลผู้ป่วยโรคเบาหวานตามฐานความรู้ (Ontology-based Clinical Reminder System for Diabetes Patient Registry) โดยมีวัตถุประสงค์เพื่อเผยแพร่ความรู้แนวทางเวชปฏิบัติ (Clinical practice guideline - CPG) ไปสู่ผู้ประยุกต์ปฏิบัติในรูปแบบที่สามารถนำไปใช้ประโยชน์ได้อย่างสะดวกรวดเร็ว และทันท่วงทีกับความต้องการ เพื่อสนับสนุนการตัดสินใจในการดูแลรักษาผู้ป่วย (Decision Support) โดยคุณสมบัติที่สำคัญของระบบคือ สามารถแจ้งเตือนความจำ (remind) ให้กับผู้ให้บริการทางการแพทย์ผ่านระบบทะเบียนผู้ป่วยเบาหวาน (diabetes patient registry) โดยแจ้งเตือนข้อมูลสำคัญ 2 ส่วน คือ แจ้งเตือนกำหนดระยะเวลา และวันที่ผู้ป่วยโรคเบาหวานแต่ละคนควรจะเข้ารับการตรวจรักษาต่างๆ เช่น ตรวจวัดระดับค่า HBA1C, ตรวจวัดระดับไขมัน, ตรวจหาภาวะแทรกซ้อนจากโรคเบาหวาน เช่น การตรวจตา, การตรวจหาระดับโปรตีนในปัสสาวะ (proteinuria) เพื่อดูการเสื่อมของไต และ การตรวจเท้า เป็นต้น โดยอิงจากข้อแนะนำในเอกสารแนวเวชปฏิบัติ และแพทย์ผู้เชี่ยวชาญ นอกจากนี้โปรแกรมยังสามารถแจ้งเตือนเมื่อผู้ป่วยมีระดับค่าผลตรวจทางห้องปฏิบัติสูงหรือต่ำกว่าค่าปกติ เช่น ค่าระดับน้ำตาล (FBS), ค่าระดับไขมัน, ค่าความดัน, ค่า HBA1C, ค่าระดับโปรตีนในปัสสาวะ เป็นต้น พร้อมทั้งข้อแนะนำในการปฏิบัติตนที่เหมาะสมให้กับผู้ป่วย โดยอิงกับข้อแนะนำจากเอกสารแนวเวชปฏิบัติ

ความท้าทายของการพัฒนาที่สำคัญจึงอยู่ที่ การออกแบบและจัดเก็บองค์ความรู้ทางด้านการดูแลรักษาโรคเบาหวานที่อิงกับเอกสารแนวเวชปฏิบัติ ในรูปแบบที่สามารถประมวลผลได้โดยโปรแกรมคอมพิวเตอร์ โดยเน้นให้สามารถนำมาแบ่งปัน และใช้ซ้ำได้ง่าย (Knowledge Sharing and Reuse) ในรูปแบบของออนโทโลยี เพื่อสะดวกต่อการนำไปใช้งานในโปรแกรมและระบบสารสนเทศทางการแพทย์ต่างๆ ได้อย่างหลากหลาย โดยระบบแจ้งเตือนความจำทางคลินิก (Clinical Reminder System) จะประยุกต์ใช้องค์ความรู้ที่ออกแบบและพัฒนาขึ้น ให้สามารถนำมาบูรณาการเชื่อมโยงกับระบบฐานข้อมูลทะเบียนผู้ป่วยโรคเบาหวานที่มีการใช้งานอยู่จริงได้ เช่น โปรแกรม DMSDD ซึ่งเป็นโปรแกรมสำหรับลงทะเบียนผลการตรวจรักษาผู้ป่วยเบาหวานที่มีใช้งานอยู่ในคลินิกโรคเบาหวานหลายแห่งทั่วประเทศ

ภาพประกอบที่ 2-14 การประยุกต์ใช้งานในโปรแกรมแจ้งเตือนความจำ (Reminde)

การประยุกต์ใช้ออนโทโลยี (Ontology application)

ออนโทโลยีถูกนำไปประยุกต์ใช้ในหลาย ๆ งาน สามารถแบ่งกลุ่มโปรแกรมที่ประยุกต์ใช้ออนโทโลยี ได้ ดังนี้

การนำออนโทโลยีไปใช้เพื่อแปลงข้อมูลให้อยู่ในรูปแบบภาษาต่าง ๆ (Neutral authoring) นำออนโทโลยีไปใช้เพื่อแปลงข้อมูลให้อยู่ในรูปแบบภาษาต่าง ๆเพื่อให้โปรแกรมอื่น ๆ สามารถใช้งานได้ประโยชน์ที่ได้จากการประยุกต์ใช้ออนโทโลยี คือ การนำความรู้มาใช้ได้อีก (Knowledge reuse)

การนำออนโทโลยีมาใช้เพื่อกำหนดรายละเอียดของซอฟต์แวร์ (Ontology as specification) ประยุกต์ใช้ออนโทโลยีเพื่อออกแบบซอฟต์แวร์ในโดเมน และรวบรวมคำศัพท์สำหรับกำหนดความต้องการในการพัฒนาซอฟต์แวร์ ประโยชน์ที่ได้จากการประยุกต์ใช้ออนโทโลยี คือ การทำคู่มือโปรแกรม การบำรุงรักษาซอฟต์แวร์ และ การนำกลับมาใช้ใหม่

การประยุกต์ใช้ออนโทโลยี เพื่อการเข้าถึงข้อมูลที่มีโครงสร้าง หรือรูปแบบต่างกัน (Common Access to Information) ออนโทโลยีจัดเตรียมคำที่สามารถเข้าใจได้ตรงกัน หรือจัดกลุ่มคำที่มีความหมายเดียวกัน ประโยชน์ที่ได้ คือ การทำงานร่วมกัน (Inter-operability) และการนำกลับมาใช้ใหม่

การประยุกต์ใช้ออนโทโลยีเพื่อการสืบค้นข้อมูลจากแหล่งข้อมูลต่าง ๆ (Ontology-based search) ประยุกต์ใช้ออนโทโลยีเพื่อการสืบค้นข้อมูลจากแหล่งข้อมูลต่าง ๆ เช่น เอกสาร เว็บเพจ หรือฐานข้อมูล แนวทางนี้ประยุกต์ใช้ออนโทโลยีในการกำหนดคอนเซพที่สอดคล้องกับคำสืบค้น

ของผู้ใช้ และใช้คอนเซพนั้นในการสืบค้นข้อมูล ทำให้ผลการสืบค้นมีความถูกต้องมากยิ่งขึ้น และเวลาที่ใช้ในการสืบค้นลดลง ซึ่งจะอธิบายรายละเอียดเพิ่มเติมในหัวข้อถัดไป

การสืบค้นข้อมูลโดยประยุกต์ใช้ออนโทโลยี

ในปัจจุบันออนโทโลยี มีความสำคัญต่อการพัฒนาแอพพลิเคชันในด้านต่างๆ ให้มีประสิทธิภาพมากขึ้น โดยเฉพาะการพัฒนาแอพพลิเคชันสำหรับสืบค้นข้อมูลที่ประยุกต์ใช้ออนโทโลยี เพื่อสืบค้นข้อมูลจากแหล่งข้อมูลต่าง ๆ เช่น เอกสาร เว็บเพจ และการสืบค้นข้อมูลในโดเมนที่เกี่ยวข้องกับวิทยาศาสตร์เพื่อชีวิต เช่น Gene Ontology Project [GO] (1999), Plant Ontology Consortium [POC] (2004)การประยุกต์ใช้ออนโทโลยีในการสืบค้นข้อมูล สามารถอธิบายได้ดังนี้ การประยุกต์ใช้ออนโทโลยีเป็นโครงสร้างพื้นฐานในการจัดการข้อมูล เช่น ประยุกต์ใช้ออนโทโลยีเป็นโครงสร้างของคอนเซพ เพื่อให้ผู้ใช้ทราบถึงขอบเขตของแหล่งข้อมูล ตัวอย่างการใช้เช่น Yahoo taxonomy ซึ่งจะแบ่งกลุ่มข้อมูลเป็น 14 กลุ่มใหญ่ที่ประกอบด้วยกลุ่มย่อยที่สอดคล้องกัน, การประยุกต์ใช้ออนโทโลยีเป็นคำศัพท์สำหรับกำหนดเมตาดาต้า (Metadata language) สำหรับสร้างดัชนี (Index) หรือสร้างแท็ก (Tag) ให้กับข้อมูล อีกแนวทางหนึ่ง คือ การประยุกต์ใช้ออนโทโลยีเพื่อแปลงคำสั่งสืบค้น (Query transformation) เช่น การประยุกต์ใช้ออนโทโลยีในการพัฒนาส่วนการติดต่อกับผู้ใช้ในการสร้าง และแก้ไขคำสั่งสืบค้น, การประยุกต์ใช้ออนโทโลยีเพื่ออ้างอิงถึงคอนเซพที่สอดคล้องกับคำสืบค้นตัวอย่างระบบการสืบค้นข้อมูลโดยใช้ออนโทโลยี ได้แก่ Knowledge-Based Discovery Tool เมื่อผู้ใช้กำหนดคำสืบค้นที่ต้องการ ระบบจะพิจารณาคำสืบค้นกับออนโทโลยี WordNet เพื่อพิจารณาหาคอนเซพที่สอดคล้องกับคำสืบค้น และให้ผู้ใช้กำหนดคอนเซพที่ถูกต้อง หลังจากนั้นคอนเซพที่มีความหมายไม่ตรงกับคำสืบค้นก็จะถูกคัดออกไป เพื่อลดความกำกวมของผลการสืบค้นในหัวข้อถัดไปจะอธิบายถึงการสืบค้นข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี ซึ่งเป็นแนวทางที่ประยุกต์ใช้ออนโทโลยี เพื่อเพิ่มประสิทธิภาพในการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ ซึ่งสืบค้นโดยพิจารณาคำสืบค้นโดยตรงให้สามารถสืบค้นโดยพิจารณาคำสืบค้นเชิงความหมายได้ (Semantic search)

การสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี

การสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี เป็นแนวทางเพิ่มประสิทธิภาพการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ให้สามารถสืบค้นข้อมูลเชิงความหมายและข้อมูลที่มีลักษณะเป็นลำดับชั้นได้ (Concept Hierarchy) เช่น การสืบค้นข้อฐานสมุนไพรในระดับตระกูล (Order)และวงศ์ (Family) โดยประยุกต์ใช้ออนโทโลยีเพื่อแปลงคำสั่งสืบค้นของผู้ใช้ให้อยู่ในรูปแบบคำสั่งสืบค้นเชิงความหมาย เช่น การพิจารณาคอนเซพในออนโทโลยีที่เฉพาะเจาะจง (Specialization) คือ การหาโหนดที่เป็นโหนดลูก (Child node) ของโหนดที่เราสนใจ, การพิจารณาคอนเซพในออนโทโลยีในระดับทั่วไป (Generalization) คือ การหาโหนดบรรพบุรุษ (Parent Node) ของโหนดที่เราพิจารณา และการพิจารณาคอนเซพข้างเคียงในออนโทโลยี (Neighborhood) ทำให้ผลการสืบค้นที่ได้มีความครบถ้วนมากยิ่งขึ้น เมื่อเทียบกับการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ทั่ว ๆ ไปที่สืบค้นโดยใช้คำสืบค้นจากผู้ใช้เพียงเท่านั้นอีกทั้งยังลดเวลาที่ใช้ในการสืบค้นอีกด้วย โดยสมมติว่า ถ้าผู้ใช้ต้องการสืบค้นคำว่า “Computer”

```
Select * from Item where Name=’Computer’
```

ผลการสืบค้นที่ได้จะมีเพียงเรคคอรด์เดียวเท่านั้น คือ RecID ที่ 123 แต่เมื่อพิจารณาออนโทโลยีกลุ่มสินค้าจะเห็นได้ว่า Computer มีความหมายเหมือนกันกับ “Calculator” และ “Data Processor” ซึ่งอธิบายด้วย ความสัมพันธ์แบบ syn-of ในออนโทโลยี นอกจากนี้ เมื่อพิจารณาความสัมพันธ์แบบ is-a จะพบว่า “PC”, “Notebook” และ “Palmtop” สอดคล้องกับคำสืบค้น ดังนั้นผลการสืบค้น คือ RecID ที่ 123,125,127, 141

ตารางที่ 2.1 ตารางผลจากการสืบค้นในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี

RecID	**Name**	**Model**	**Price**
129	Monitor	ELSA	1000$
135	Keyboard	ITT	80$
136	Desktop	IBM	1000$
140	MacPC	MAC	2000S
141	Calculator	SIEMEN	1500S

การพัฒนาออนโทโลยีเพื่อเพิ่มประสิทธิภาพการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ ต้องพิจารณาถึงแนวทางการจัดเก็บและสืบค้นออนโทโลยี รวมถึงวิธีการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ที่สอดคล้องกับออนโทโลยี

วิธีการสืบค้นสารสนเทศบนเว็บ

วิธีการสืบค้นสารสนเทศบนเว็บแบ่งออกได้เป็น 3 ประเภท ดังนี้

1. เสิร์จเอ็นจิน (Search Engine) เป็นระบบซอฟท์แวร์ที่มีซอฟต์แวร์หุ่นยนต์ (Robot Software) ซึ่งบางครั้งเรียกว่า Spider หรือ Crawler ทำหน้าที่ท่องไปในเว็บไซต์ต่างๆ ในอินเตอร์เน็ต เพื่อรวบรวมเอกสารบนเว็บ เพื่อนำมาสร้างเป็นฐานดัชนีสำหรับการสืบค้นเอกสารบนเว็บ เพื่อทำให้การสืบค้นกระทำได้อย่างรวดเร็ว ดังนั้นกระบวนการจัดเตรียมดัชนี จึงเป็นขั้นตอนสำคัญอย่างยิ่งสำหรับการสืบค้นข้อมูลบนเว็บ เนื่องจากการเตรียมข้อมูลหรือดัชนีสำหรับการสืบค้นเป็นกระบวนการที่จัดเตรียมโดยใช้ซอฟต์แวร์ ทำให้สามารถอัพเดตข้อมูลได้ตลอดเวลา นั่นคือสามารถส่งซอฟต์แวร์หุ่นยนต์ออกไปรวบรวมเอกสารบนเว็บได้ตลอดเวลา ตัวอย่างของเสิร์จเอ็นจินที่มีอยู่ในปัจจุบันและเป็นที่นิยมคือ Google อย่างไรก็ตาม วิธีการจัดทำดัชนีด้วยวิธีการนี้ ความถูกต้องของการสืบค้นจะขึ้นกับฐานข้อมูลดัชนี และวิธีการสืบค้น

2. ไดเร็กทอรี (Directory) แตกต่างจากเสิร์จเอ็นจินตรงที่ไดเร็กทอรีมีการแบ่งแยกหมวดหมู่ของเอกสารโดยคน เว็บไซต์ใดที่ต้องการมีรายชื่อในไดเร็กทอรีต้องติดต่อผู้ดูแลไดเร็กทอรี เพื่อให้ผู้ดูแลไดเร็กทอรีจำแนกแยกประเภทของเอกสารบนเว็บนั้นๆให้อยู่ในหมวดหมู่ที่เหมาะสม เนื่องจากไดเร็กทอรี ถูกจัดเป็นหมวดหมู่โดยมนุษย์ ซึ่งจะมีการอ่านและพิจารณาแยกหมวดหมู่ได้ถูกต้องกว่าคอมพิวเตอร์ การค้นหาเอกสารบนเว็บด้วยไดเร็กทอรีจึงให้ผลลัพธ์ที่ตรงกับความต้องการมากกว่าเสิร์จเอ็นจิน แต่ด้วยเหตุผลไดเร็กทอรีมีการจัดแบ่งหมวดหมู่โดยคน ทำให้จำนวน

เอกสารของไดเร็กทอรีมีน้อยกว่าเอกสารที่มีอยู่ในเสิร์จเอ็นจิน ตัวอย่าง ไดเร็กทอรีที่มีอยู่ในปัจจุบัน คือ Yahoo

3. ไฮบริดเสิร์จเอ็นจิน (Hybrid Search Engine) เป็นเสิร์จเอ็นจินที่เป็นการผสมผสานการทำงานของเสิร์จเอ็นจินและไดเร็คทอรี่เข้าด้วยกัน มีข้อดีของทั้งเสิร์จเอ็นจินและไดเร็คทอรีรวมอยู่ด้วยกัน

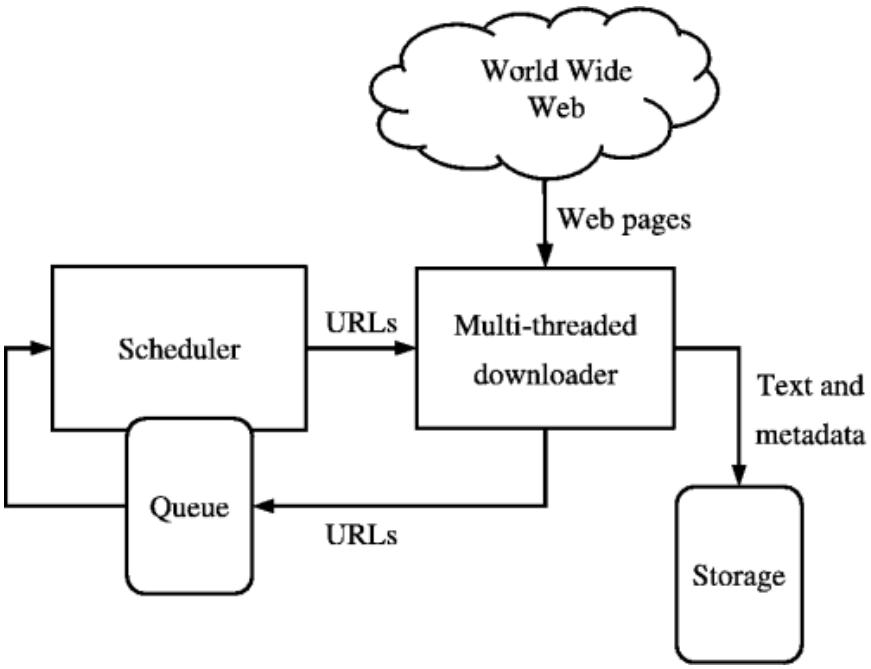
Search Engine

Search Engine คือ เครื่องมือการค้นหาข้อมูลผ่านอินเตอร์เน็ต ที่ทุกคนสามารถเข้าไปค้นหาข้อมูลในอินเตอร์เน็ตได้อย่างง่ายดาย โดยกรอกข้อมูลที่ต้องการค้นหา หรือ Keyword เข้าไปที่ช่อง Search Box เท่านั้น โดยข้อมูลที่เราต้องการค้นหาจะถูกแสดงออกมาเป็นจำนวนมากเพื่อให้เราเลือกข้อมูลที่เราสนใจมากที่สุดเอามาใช้งาน ลักษณะของการแสดงผลของ Search Engine นั้นจะทำการแสดงผลแบบเรียงอันดับ Search Result ผ่านหน้าจอคอมพิวเตอร์ โดย Search Engine สามารถแบ่งการหลักการทำงานที่แตกต่างกันและการจัดอันดับข้อมูลออกได้เป็น 3 ประเภท คือ

1) Crawler Based Search Engines เป็นเครื่องมือการค้นหาบนอินเตอร์เน็ตแบบอาศัยการบันทึกข้อมูล และจัดเก็บข้อมูลเป็นหลัก ซึ่งจะเป็นจำพวก Search Engine ที่ได้รับความนิยมสูงสุด เนื่องจากให้ผลการค้นหาที่แม่นยำที่สุด และการประมวลผลการค้นหาสามารถทำได้อย่างรวดเร็วจึงทำให้มีบทบาทในการค้นหาข้อมูลมากที่สุดในปัจจุบันมีการทำงานดังในรูปที่ 13 โดยมีองค์ประกอบหลักอยู่ 2 ส่วน คือ

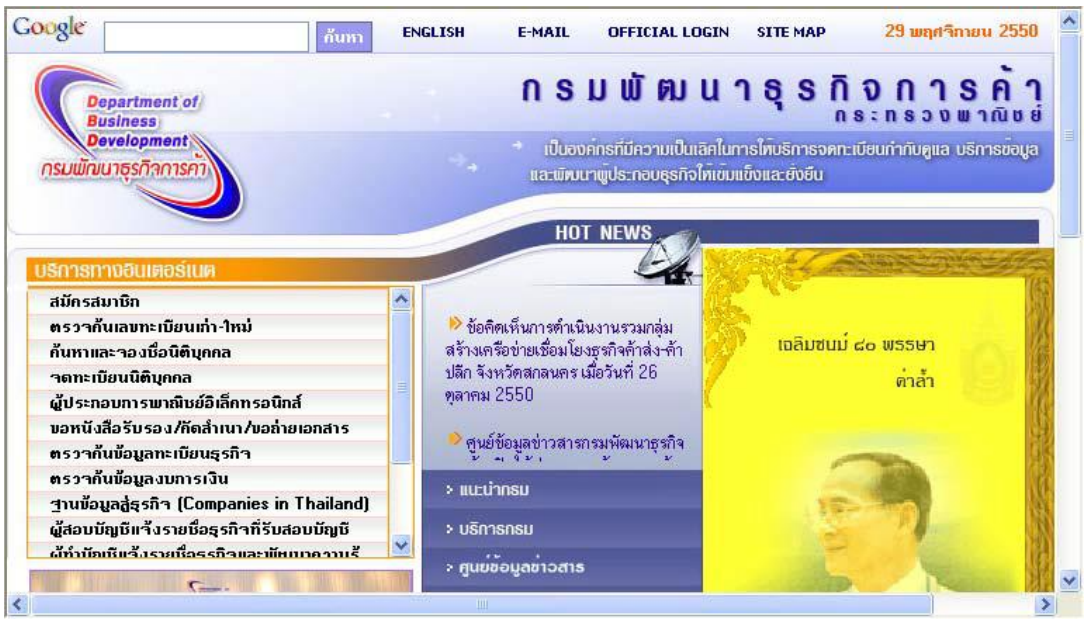
- ฐานข้อมูล โดยส่วนใหญ่แล้ว Crawler Based Search Engine จะมีฐานข้อมูลเป็นของตัวเอง ที่มีระบบการประมวลผล และจัดอันดับที่เฉพาะเป็นเอกลักษณ์ของตนเอง

- Software คือ เครื่องมือหลักสำคัญที่สุดอีกส่วนหนึ่งสำหรับ Search Engine ประเภทนี้ เนื่องจากต้องอาศัยโปรแกรมเล็กๆ ทำหน้าที่ในการตรวจหา และทำการจัดเก็บข้อมูล หน้าเพจหรือเว็บไซด์ต่างๆ ในรูปแบบของการทำสำเนาข้อมูล เหมือนกับต้นฉบับทุกอย่าง ซึ่งเราจะรู้จักกันในนาม Spider หรือ Web Crawler หรือ Search Engine Robots



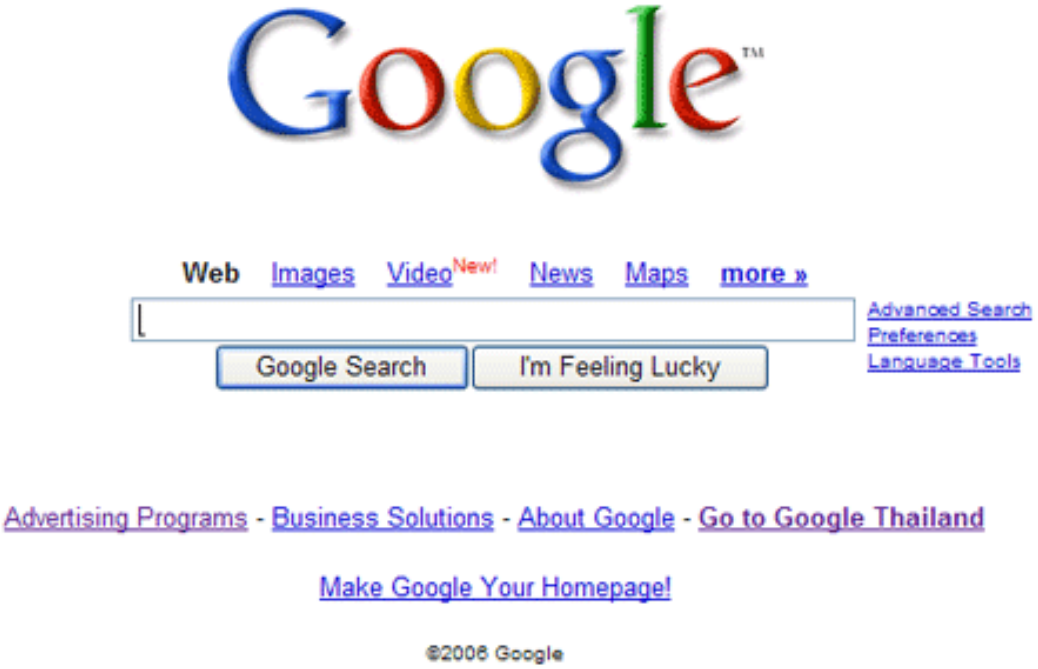
ภาพประกอบที่ 2-15 ลักษณะการทำงานของ Web Crawler

การทำ Web Crawler แบบเฉพาะเจาะจงนั้น มักจะไม่ได้เน้นในส่วนของการทำ Multi-Thread เพื่อแบ่งย่อยหลายๆ งาน แล้วแบ่งๆ แยกย้ายกันไปอ่านตามแต่ละลิงค์ แต่อยู่ตรงที่ตัว Web Crawler เองนั้น ต้องรู้จักการ Login เข้าไปสู่ระบบดังกล่าว แล้วจึงเข้าไปยังหน้าเว็บไซด์ลึกๆ เพื่ออ่านข้อมูล จากนั้นจึงเจาะจงเลือกเฉพาะบางลิงค์ บางตำแหน่งเท่านั้นเพื่อกระโดดต่อไป นอกจากนี้การออกแบบโครงสร้างข้อมูลใน Virtual Memory ให้สอดรับกับ Tag HTML ทั้งปวงของเว็บไซด์นั้นๆ อีกทั้งต้องสามารถกำหนดจุดเริ่มต้น หรือสิ้นสุดในแต่ละบล็อคที่ต้องการอ่านลิงค์อีกด้วยและหากการเขียน Software ดังกล่าวไม่เขียนเป็นระบบเปิด ก็จะใช้ได้เฉพาะกับเว็บไซด์ใดเว็บไซด์หนึ่งเท่านั้น ดังนั้นการทำ Web Crawler แบบเฉพาะเจาะจง จึงหวังที่ผลลัพธ์ที่ได้จากเว็บไซด์นั้นเป็นสำคัญ ตัวอย่างเว็บไซด์ที่เป็นระบบปิด เมื่อต้องการค้นหาข้อมูล จำเป็นจะต้องล็อกอินเข้าสู่ระบบก่อนทุกครั้ง ดังภาพประกอบที่ 2-16



ภาพประกอบที่ 2-16 ลักษณะการทำงานของ Web Crawler

ตัวอย่างเว็บที่ใช้เทคนิคนี้ เช่น Google, Yahoo, Live, Search, Technorati เป็นต้น ซึ่งลักษณะการทำงานและการเก็บข้อมูลของ Web Crawler แต่ละแห่งจะมีวิธีการเก็บและการจัดอันดับข้อมูลที่แตกต่างกันดังภาพประกอบที่ 2-17



ภาพประกอบที่ 2-17 ตัวอย่างเว็บไซต์ที่ทำงานแบบ Web Crawler

2) Web Directory หรือ Blog Directory เป็นสารบัญเว็บไซต์ที่สามารถค้นหาข่าวสารข้อมูลด้วยหมวดหมู่ข่าวสารข้อมูลที่เกี่ยวข้องกัน ในปริมาณมากๆ คล้ายๆ กับสมุดหน้าเหลือง ซึ่งจะมีการสร้างดรรชนี มีการระบุหมวดหมู่อย่างชัดเจน ซึ่งจะช่วยให้การค้นหาข้อมูลต่างๆ ตามหมวดหมู่นั้นๆ ได้รับการเปรียบเทียบอ้างอิง เพื่อหาข้อเท็จจริงได้ ในขณะที่เราค้นหาข้อมูล เพราะจะมีเว็บไซต์มากมาย หรือ Blog มากมายที่มีเนื้อหาคล้ายๆ กัน ในหมวดหมู่เดียวกันให้เราเลือกที่จะหาข้อมูลได้อย่างตรงประเด็นที่สุด ซึ่งจะลดระยะเวลาในการค้นหาได้มาก ดังภาพประกอบที่ 2-18

ภาพประกอบที่ 2-18 ตัวอย่าง Web Directory ที่มี Search Engine มากมายใช้เป็นฐานข้อมูล

- OPD หรือ Dmoz ที่เป็นที่นิยมเป็น Web Directory ที่ใหญ่ที่สุดในโลก Search Engine หลายๆ แห่ง มักใช้ข้อมูลจากแห่งนี้เกือบทั้งสิ้น เช่น Google, AOL, Yahoo, Netscape และอื่นๆ อีกมากมาย ODP มีการบันทึกข้อมูลประมาณ 80 ภาษาทั่วโลก (URL: http://www.dmoz.org)

- สารบัญเว็ปไทย Sanook ก็เป็น Web Directory ที่มีชื่อเสียงอีกแห่งหนึ่ง และเป็นที่รู้จักกันมากในเมืองไทย (URL: http://webindex.sanook.com)

- Blog Directory อย่าง BlogFlux Directory ที่มีการเก็บข้อมูลเกี่ยวกับบล็อกมากมายตามหมวดหมู่ต่าง ๆ

Meta Search Engine

เป็น Search Engine ที่ใช้หลักการในการค้นหาโดยอาศัย Meta Tag ในภาษา HTML ซึ่งจะมีการประกาศชุดคำสั่งต่าง ๆ เป็นรูปแบบของ Text Editor ด้วยภาษา HTML คือ ชื่อผู้พัฒนา คำค้นหา เจ้าของเว็บหรือบล็อก คำอธิบายเว็บ หรือบล็อกอย่างย่อ

ผลการค้นหาของ Meta Search Engine นี้มักไม่แม่นยำอย่างที่คิด เนื่องจากบางครั้งผู้ให้บริการหรือผู้ออกแบบเว็บสามารถใส่อะไรเข้าไปก็ได้มากมายเพื่อให้เกิดการค้นหาและพบเว็บหรือ บล็อกของตนเองและ อีกประการหนึ่งก็คือ มีการอาศัย Search Engine Index Server หลาย ๆ แห่งมาประมวลผลรวมกัน จึงทำให้ผลการค้นหาข้อมูลต่าง ๆ ไม่เที่ยงตรงเท่าที่ควร ตัวอย่างคำสั่งของ Tag META จะมีการประกาศรายละเอียดของเว็บไซด์นั้นๆ ภาพประกอบที่ 2-19

```
<HEAD>
<TITLE>Stamp Collecting World</TITLE>
<META name="description" content="Everything you wanted to know
about stamps, from prices to history.">
<META name="keywords" content="stamps, stamp collecting,
stamp history, prices, stamps for sale">
</HEAD>
```

ภาพประกอบที่ 2-19 ตัวอย่างคำสั่งใน Tag META เพื่อประกาศข้อมูลของเว็บ

XML (eXtensible Markup Language)

ภาษา XML (eXtensible Markup Language) เป็นส่วนหนึ่งของ SGML (the Standard Generalized Markup Language) โดยพัฒนาให้สามารถใช้งานได้ง่ายกว่าและมีความยืดหยุ่นมากกว่า โดย XML ถูกนำเสนอในปี ค.ศ.1996 ภายใต้การดูแลของ W3C (World Wide Web Consortium) ภาษา XML เป็น Meta-Language ที่ใช้แท็กในการสื่อความหมายของข้อมูล ซึ่งผู้พัฒนาสามารถกำหนดแท็กขึ้นใช้งานได้เองตามความต้องการเริ่มแรกภาษา XML ถูกพัฒนาขึ้นมาเพื่อแก้ไขปัญหาของภาษา HTML แต่ต่อมาได้เพิ่มขอบเขตความสามารถออกไป ได้แก่ สามารถแลกเปลี่ยนข้อมูลกันได้โดยไม่ขึ้นกับ Platform ใดPlatform หนึ่ง ทำให้คอมพิวเตอร์สามารถติดต่อสื่อสารเพื่อแลกเปลี่ยนข้อมูลได้อย่างไร้ขีดจำกัดและต้องการให้ภาษา XML เป็นภาษามาตรฐานของเว็บที่ใช้กันทั่วโลกเพื่อให้ผู้พัฒนาที่เป็นรูปแบบเดียวกัน

RDF (Resource Description Language)

สิ่งที่สำคัญของแนวความคิดเว็บเชิงความหมาย (Semantic Web) คือการทำให้เครื่องคอมพิวเตอร์สามารถเข้าใจความหมายของข้อมูลได้

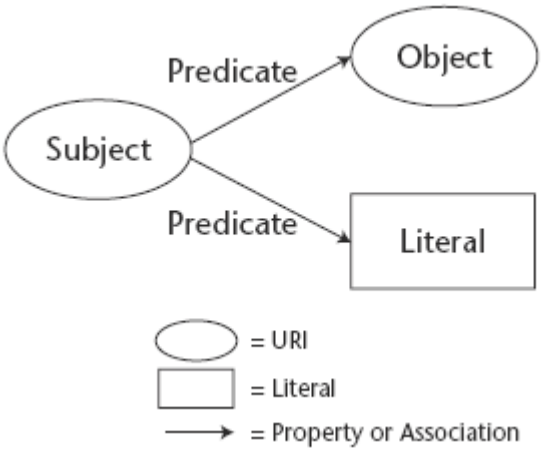
RDF เป็นมาตรฐานโครงสร้างการอธิบาย Metadata บนเว็บ ที่ถูกพัฒนาขึ้นมาโดยองค์กร W3C (World Wide Web Consortium) ซึ่งมีจุดมุ่งหมายที่จะพัฒนา RDF ให้เป็นส่วนสำคัญของ Semantic Web และเพื่อง่ายต่อการใช้งานในระบบที่ต่าง Platform กัน RDF จึงเขียนให้อยู่ในรูปแบบโครงสร้างของภาษา XML (World Wide Web Consortium, 2002) ลาสสิร่าและสวิก (Lassila & Swick, 1999) ได้ทำการวิจัยโดยกล่าวว่า RDF สามารถนำไปใช้ประโยชน์ในหลากหลายด้าน ยกตัวอย่างเช่น การสืบค้นเอกสารบนเว็บ ด้วย Search Engine ได้อย่างมีประสิทธิภาพมากขึ้น, สามารถอธิบายข้อมูลและความสัมพันธ์ระหว่างข้อมูลโดยเฉพาะเว็บไซต์, เว็บเพจ โดย Intelligent Software Agent ซึ่งทำให้ง่ายต่อการแบ่งปันข้อมูลและการแลกเปลี่ยนข้อมูลซึ่งกันและกัน

แนวความคิดพื้นฐานของ RDF ประกอบไปด้วย 3 ส่วนประกอบไปด้วย Resources, Properties, Statements ซึ่งแต่ละส่วนสามารถอธิบายแต่ละส่วนได้ดังนี้

1. Resources คือ สิ่งใดๆ ที่ต้องการจะอธิบาย โดยที่ทุกๆ Resource จะต้องถูกระบุด้วย URI (Universal Resource Identifier) ซึ่ง URI ที่ว่านี้อาจจะเป็น URL (Web Address) หรือสิ่งที่ระบุให้ Resource นั้นจะต้องมีคุณสมบัติที่ไม่ซ้ำกัน

2. Properties คือ การอธิบายคุณสมบัติหรือความสัมพันธ์ระหว่าง Resource ซึ่ง Properties จะต้องถูกระบุด้วย URI เช่นกัน

3. Statements คือ RDF จะมีการอธิบายข้อมูลในรูปแบบที่เรียกว่า “Triple” ประกอบด้วย Subject, Predicate และ Object



ภาพประกอบที่ 2-20 ส่วนประกอบของ Triple

จากภาพประกอบที่ 2-20 แสดงส่วนประกอบของ Triple ประกอบไปด้วย

1) Subject คือ ทรัพยากรที่อธิบาย แทนที่ด้วยสัญลักษณ์วงรี ซึ่งแทนด้วย URI

2) Predicate คือ คุณสมบัติของทรัพยากรแทนด้วยลูกศรชี้จากส่วน Subject ชี้ไปยังส่วนของ Object หรือ Literal

3) Object คือ ค่าของคุณสมบัติ แทนด้วยสัญลักษณ์ที่เป็นวงรี ซึ่งแทนที่ด้วย URI แต่หากเป็นค่าของอักขระจะเรียกว่า Literal ซึ่งจะมีลักษณะเป็นสัญลักษณ์รูปสี่เหลี่ยม

RDFS (RDF Schema)

RDF Schema เป็นส่วนต่อขยายจาก RDF ใช้ในการอธิบายโครงสร้างของ Metadata ของทรัพยากรบนเว็บโดยกำหนด Vocabulary เพื่ออธิบายโครงสร้างของ Metadata ซึ่งประกอบไปด้วย Property และค่าของ Property โดยอธิบายให้อยู่ในรูปของ Class และ Property โดยสามารถสืบทอดเป็น SubClass และ SubProperty ได้ เมื่อเอกสาร RDF มี RDF Schema กำกับจะช่วยให้เครื่องอนุมานข้อมูลได้ (Inference Engine)

RDF Schema เป็นส่วนที่ขยายต่อมาจาก RDF โดยสามารถอธิบายกลุ่มของทรัพยากรบนเว็บที่มีความเกี่ยวข้องกันและสามารถเชื่อมต่อความสัมพันธ์ระหว่างทรัพยากรบนเว็บ ยกตัวอย่าง เช่น “Pupil” เป็นกลุ่มประเภทของ “Student” และ “Student” ก็เป็นคลาสย่อยของ “People” เป็นต้น (Swartz, internet, 2002)เป็นการแสดงความสัมพันธ์ระหว่างคลาสต่างๆ โดยความสัมพันธ์ตัวอย่างเช่น ระบุว่า สุนัข (Dog), แมว (Cat), และลา (Donkey) เป็นซับคลาสของ สัตว์ (Animal) ส่วน แมวเปอร์เซีย (Persian Cat) มีชนิดเป็นแมว

ดังนั้น เมื่ออธิบายโดยใช้ RDFS ไว้เครื่องคอมพิวเตอร์ก็สามารถเข้าใจได้ว่า แมวเปอร์เซียอยู่ในชนิดของสัตว์

RDQL (RDF Data Query Language)

RDQL (RDF Data Query Language) เป็นภาษาในการสืบค้นข้อมูลบนโครงสร้างของ RDF ซึ่งคล้ายกับการดำเนินการของภาษาในการสอบถามข้อมูลแบบโครงสร้าง SQL ของฐานข้อมูลเชิงสัมพันธ์ โดยเมื่อรับคำสั่งมาจากโปรแกรมประยุกต์แล้ว RDQL จะดำเนินการสืบค้นข้อมูลพร้อมกับส่งผลลัพธ์ที่ตรงกับเงื่อนไขกลับมายังผู้ใช้งาน (Seaborne, internet, 2004)

OWL (Web Ontology Language)

ถึงแม้ว่า RDF Schema เป็นภาษาที่สามารถอธิบายออนโทโลยี แต่เนื่องจากภาษา RDF/RDFS ยังมีข้อจำกัดที่สามารถสรุปได้ดังนี้

Local scope ของ properties เนื่องจากคำสั่ง rdf:range ใช้ในการกำหนด Range ของ Property ให้กับทุกคลาส เช่น Property “การกิน” มีโดเมนคือ ‘สัตว์’ และมี Range คือ ‘เนื้อสัตว์’ ดังนั้นคลาสที่เป็นคลาสย่อยของ ‘สัตว์’ จะต้องมีข้อกำหนดของ Property ตามนี้ ไม่สามารถกำหนดได้คลาส ‘วัว’ ซึ่งเป็นคลาสย่อยของสัตว์มี Range ของ Property ‘การกิน’ เป็น ‘พืช’ Disjoint ของคลาส ตัวอย่างเช่น คลาส ‘ผู้ชาย’ และ ‘ผู้หญิง’ เป็น Disjoint class กัน หมายถึงไม่มี instance ใดที่เป็นสมาชิกของทั้งคลาส “ผู้ชาย” และ “ผู้หญิง”

Boolean combination ของคลาส ในบางครั้งเราต้องการที่จะสร้างคลาสใหม่จากคลาสเดิมโดยการนำคลาสเดิมมา union, intersection หรือ complement กัน เช่น คลาส ‘คน’ เกิดจากการ union กัน ของคลาส ‘ผู้ชาย’ และ ‘ผู้หญิง’ ซึ่งไม่สามารถกำหนดได้ใน RDFS Cardinality restriction การสร้างข้อกำหนดของคุณสมบัติให้มีค่าเป็นตัวเลข เช่น คลาส Person จะมีแม่ (Mother) ได้ 1 คนเท่านั้น ซึ่งไม่สามารถกำหนดได้ใน RDFS

ดังนั้นทางองค์กร W3C (World Wide Web Consortium) จึงได้ทำการพัฒนา OWL (Web Ontology Language) เพื่อเป็นส่วนขยายต่อจากภาษา RDF/RDFS ภาษา OWL จัดเป็นองค์ประกอบหนึ่งในงานเว็บเชิงความหมาย (Semantic Web) ที่ใช้ในการบรรยายข้อมูลเชิงความหมาย สามารถกำหนดโครงสร้างข้อมูลในลักษณะลำดับชั้น และอธิบายข้อมูล (Metadata) ที่มีความสัมพันธ์ในระบบฐานข้อมูลได้ รวมทั้งสามารถรองรับการบรรยายข้อมูลเชิงตรรกะ ชนิดข้อมูล และตัวบ่งปริมาณได้ ทำให้ข้อมูลที่ถูกแทนที่นั้นมีความหมายมากยิ่งขึ้น ลักษณะการบรรยายจะอยู่ในรูปของคลาส คุณสมบัติของคลาส และความสัมพันธ์ของคลาส เพื่ออธิบายเอ็นทิตี้ (Entity) และความสัมพันธ์ต่างๆ ที่เกิดขึ้น ซึ่งทางองค์กร W3C ได้แบ่งภาษา OWL ออกเป็น 3 กลุ่มคือ

- OWL Lite เหมาะสำหรับการเก็บข้อมูลความสัมพันธ์แบบระดับชั้น (Hierarchy) และมีข้อกำหนดอย่างง่าย

- OWL DL (Description Logic) เป็นภาษาที่รองรับคำสั่งทุกคำสั่งของภาษา OWL แต่จะต้องเขียนภายใต้ข้อกำหนด เช่น คลาสสามารถเป็นซับคลาสของคลาสอื่นๆ ได้มากว่าหนึ่งคลาส

- OWL Full เป็นภาษาที่มีความซับซ้อนมากที่สุด โดยมีลักษณะพิเศษ เช่น Syntax อิสระจาก RDF, คลาสสามารถเป็น Instance ได้ทั้ง OWL DL และ OWL Full ต่างก็สนับสนุนเซตของภาษา OWL แต่มีข้อจำกัดของคุณลักษณะบางอย่างที่แตกต่างกันบนพื้นฐานของ RDFS โดย OWL Full จะมีการผสมผสานระหว่าง OWL และ RDFS โดยไม่มีการบังคับในส่วนการแบ่งคลาส การ

กำหนดคุณสมบัติ และค่าของข้อมูล ส่วน OWL DL จะมีข้อบังคับในการใช้ RDF การกำหนดคลาส การกำหนดคุณสมบัติและค่าของข้อมูล เป็นต้น

โครงสร้างของภาษา OWL (Ontology Web Language)

1) Namespace

การกำหนด Namespace จะประกาศไว้ที่ส่วนเริ่มต้นของเอกสาร เพื่อเป็นการกำหนดกลุ่มในการอ้างอิงข้อมูลเอกสาร OWL ที่ถูกสร้างจะขึ้นอยู่กับโครงสร้างที่ถูกนิยามด้วย RDF/RDFS และชนิดข้อมูลของXML Schema การเขียน Namespace จะประกาศไว้ภายใต้คำสั่ง rdf:RDF syntax ตัวอย่างในภาพประกอบที่ 2-21 แสดงการกำหนด Namespace

```
xmlns=”http://www.owl-ontologies.com/car.owl#”
xmlns:rdf=”http://www.w3.org/1999/02/22-rdf-syntax-ns#
xmlns:rdfs=”http://www.w3.org/2000/01/rdf-schema#”
xmlns:owl=”http://www.w3.org/2002/07/owl#”
```

ภาพประกอบที่ 2-21 ตัวอย่างการกำหนด Namespace

2) Ontology Headers

แสดงการอธิบายรายละเอียดเบื้องต้นของออนโทโลยี ว่าเป็นโครงสร้างข้อมูลเกี่ยวกับอะไร ภายใต้ Element <owl:Ontology rdf:about=””> ประกอบด้วยคำสั่ง <owl:versionInfo> ใช้แสดงรุ่นของข้อมูลที่สร้าง คำสั่ง <rdf:comment> ใช้แสดงการอ้างอิงเกี่ยวกับแหล่งข้อมูลว่ามาจากที่ใด ดังอย่างการแทนค่าใช้งานแสดงดังภาพประกอบที่ 2-22

```
<owl:Ontology rdf:about=”MyCarOntology”/>
<owl:Class rdf:ID=”Car”/>
```

ภาพประกอบที่ 2-22 ตัวอย่าง Ontology Header

3) การกำหนดคลาส

ในการอธิบายคลาสในออนโทโลยี จะมีคลาสเริ่มต้นคือ owl:Class โดยกำหนดให้ owl:Class เป็นคลาสใหญ่ที่สามารถครอบคลุม ทุกคลาสข้อมูลได้ ดังนั้นไม่ว่าผู้ใช้งานกลุ่มใดสร้างคลาสขึ้นมา จะเสมือนว่าเป็นสมาชิกอยู่ภายใต้คลาส owl:class

กำหนดคลาส “Island” เป็นคลาสย่อย (Subclass) ของคลาส “Natural Place” ความสัมพันธ์แบบคลาสย่อยนี้ทำให้เกิดคลาสทั่วไป (Generic class) และคลาสจำเพาะเจาะจง (Specific class) กล่าวคือ Natural Place เป็นคลาสที่แสดงความหมายของสถานที่ท่องเที่ยวเชิงธรรมชาติโดยทั่วไป ในขณะที่ Island คือคลาสที่แสดงความหมายที่จำเพาะเจาะจงว่าเป็นสถานที่ท่องเที่ยวที่เป็นเกาะ ซึ่งจากความสัมพันธ์ดังกล่าวนี้สามารถเขียนอธิบายด้วยภาษา OWL ได้ดังภาพประกอบที่ 2-23

```
<owl:Class rdf:ID=”NaturalPlace”/>
<owl:Class rdf:ID=”Island”/>
<rdfs:subClassOf
rdf:resource=”#NaturalPlace”/>
```

ภาพประกอบที่ 2-23 ตัวอย่างการกำหนดคลาสและคลาสย่อย

4) การกำหนดคุณสมบัติ (Property)

การกำหนดคุณสมบัติของคลาสใน OWL สามารถกำหนดได้ 2 ประเภทคือ

- การกำหนดคุณสมบัติด้วย owl:DatatypeProperty เพื่อกำหนดการอธิบายคุณสมบัติของคลาสที่เป็นค่าชนิดพื้นฐาน เช่น การอธิบายข้อมูลราคาสินค้า

- การกำหนดคุณสมบัติด้วย owl:ObjectProperty เพื่อกำหนดการอธิบายข้อมูลซึ่งต้องการอธิบายคุณสมบัติของคลาส ซึ่งเป็น Resource หรือการกำหนดการเชื่อมโยงความสัมพันธ์ระหว่างคลาส 2 คลาส

5.) การกำหนดความสัมพันธ์ระหว่าง Resource

ทรัพยากร (Resource) หมายถึง คลาสที่มีการอ้างอิงได้ด้วยการกำหนด URL ซึ่งการกำหนดความสัมพันธ์ระหว่างทรัพยากรเป็นการกำหนดเงื่อนไข Domain และ Range ให้กับการอธิบายคุณสมบัติที่เชื่อมโยงระหว่าง Resource ดังตัวอย่างในภาพประกอบที่ 2-11 เป็นการอธิบายข้อมูลราคาด้วยคุณสมบัติ hasPrice ซึ่งสามารถกำหนดโดเมนคือ “Car” และ Range คือชนิดพื้นฐานซึ่งเป็นตัวเลขจำนวนเต็ม (Integer)

```
<owl:DatatypeProperty rdf:ID=”hasPrice”>
<rdfs:domain rdf:resource=”#Car”/>
<rdfs:range
rdf:resource=”http://www.w3.org/2001/XMLSchema#int”/>
</owl:DatatypeProperty>
```

ภาพประกอบที่ 2-24 การกำหนดความสัมพันธ์ระหว่างทรัพยากร (Resource)

Lexical Chains

Lexical Chains เป็นวิธีการจัดหมวดหมู่ของคำที่มีความหมายอยู่ในกลุ่มเดียวกันจากประโยค บทความหรือวารสาร (Morris, & Hirst, 1991) เพื่อใช้ในการจับใจความสำคัญของบทความและหาโครงสร้างของบทความการสร้าง Lexical Chains นั้นมีอยู่ 2 ขั้นตอนหลักคือ

ขั้นตอนการหาคำสำคัญ (Candidate Word) และขั้นตอนการสร้าง Lexical Chains ขั้นตอนการหาคำสำคัญเป็นการเลือกคำสำคัญจากบทความซึ่งจะมีวิธีการเลือกคำสำคัญโดยการเลือกเอาเฉพาะคำที่สื่อความหมายได้ซึ่งได้แก่ คำนาม (Noun) ต่าง ๆ และจะไม่ในใจคำประเภทอื่น ๆ เช่น คำสันธาน (Pronouns) คำสรรพนาม (Preposition) คำศัพท์ (Adjective) คำกริยาวิเศษ (Adverb) คำกริยาต่าง ๆ และคำศัพท์ที่ใช้บ่อย ๆ ในประโยค แต่สื่อความหมายไม่ได้ขั้นตอนการสร้าง Lexical Chains เป็นการนำคำสำคัญที่ได้จากขั้นตอนการเลือกคำสำคัญมาจัดกลุ่มโดยอาศัยฐานข้อมูล Thesaurus และฐานข้อมูล Word Net ที่มีการแยกประเภทและจัดกลุ่มคำศัพท์เป็นหมวดหมู่

ฐานความรู้ธีซอรัส (Thesaurus)

ธีซอรัส (นฤมล ปราชญโยธิน, ทวีศักดิ์ กออนันตกูล, เปรมิน จินดาวิมลเลิศ, 2536) เป็นคลังคำ แต่เป็นคลังคำของศัพท์ควบคุม คือ จะกำหนดหรือแนะให้ใช้ศัพท์คำใดคำหนึ่ง เป็นตัวแทนของกลุ่มคำที่มีรูปต่างกัน หรือเขียนต่างกันแต่มีความหมายเดียวกัน และจะประมวลศัพท์ที่อยู่ในกลุ่มเดียวกันไว้ด้วยกัน โดยมีการแสดงโครงสร้างความสัมพันธ์ของคำในกลุ่ม และมีสัญลักษณ์แสดงความสัมพันธ์กำกับไว้ภายใต้รูปแบบที่ได้มาตรฐาน โดยสัญลักษณ์ที่ใช้แทนความสัมพันธ์จะมีรูปแบบเฉพาะตัวมากกว่าบัญชีคำศัพท์ประเภทอื่น ๆ

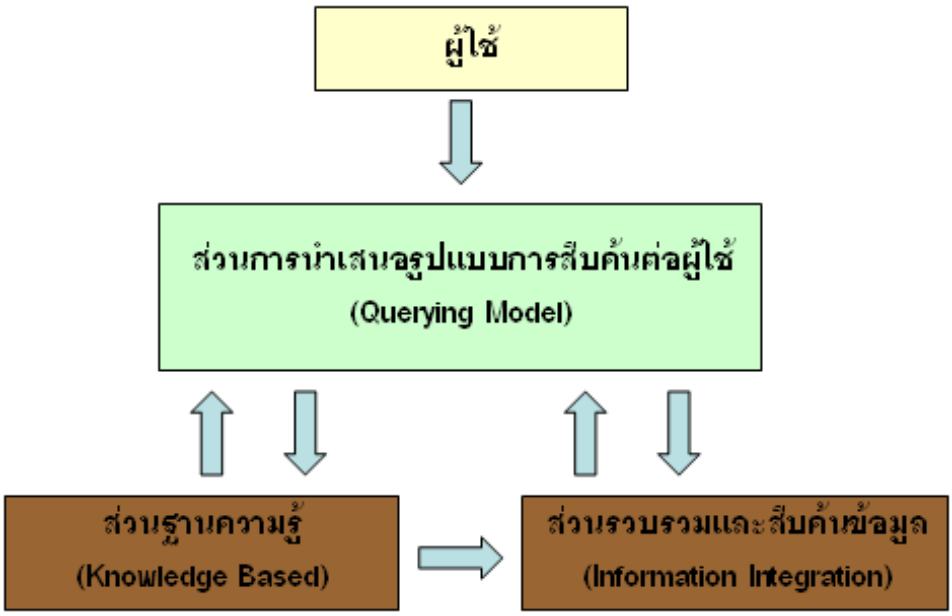
วัตถุประสงค์ของธิซอรัส คือ จัดทำขึ้นเพอื่ เป็นเครื่องมือที่มีประสิทธิภาพในการจัดเก็บและค้นคืนสารสนเทศด้วยวิธีการเข้าถึงสารสนเทศผ่านทางเนื้อหา (Subject access) คำว่าประสิทธิภาพในที่นี้ได้แก่ความสามารถในการสืบค้นข้อมูลเรื่องใดเรื่องหนึ่งได้อย่างครบถ้วน โดยไม่ต้องทราบคำศัพท์ที่เกี่ยวข้องทั้งหมดและความสามารถในการช่วยเหลือให้ผู้จัดทำดรรชนีและผู้สืบค้นใช้ศัพท์ที่เป็นมาตรฐานเดียวกัน ทำให้การกำหนดคำแทนสาระของเอกสารทีความเที่ยงตรง

ฐานความรู้เวิร์ดเนต

ฐานความรู้เวิร์ดเนตหรือฐานความรู้อภิธานเวิร์ดเนตเป็นฐานข้อมูลคำศัพท์ภาษาอังกฤษโดยการแบ่งการเก็บข้อมูลคำศัพท์ตามชนิดของคำ เช่น คำนาม คำกริยา คำวิเศษณ์ เป็นต้น การจัดโครงสร้างคำศัพท์ในฐานข้อมูลเวิร์ดเนตจะจัดตามความหมายของคำ (Word meaning) มากกว่ารูปของคำ (Word Form) โดยที่คำหนึ่งอาจมีได้หลายความหมายและคำหลายคำอาจมีความเดียวกัน (Chirstiane, 1998)

ระบบสืบค้นข้อมูล

จากการศึกษาการสืบค้นข้อมูลผ่านทางระบบเครือข่าย พบว่า ข้อมูลมีการกระจัดกระจายกันอยู่ในฐานข้อมูลต่างๆ ซึ่งยุ่งยากและเสียเวลาในการสืบค้น ดังนั้นจึงได้จัดการแบ่งระบบสืบค้นออกเป็น 3 ส่วน ได้แก่ ฐานความรู้ (Knowledge Based) การรวบรวมและสืบค้นข้อมูล (Information Integration) และการนำเสนอรูปแบบการสืบค้นต่อผู้ใช้ (Querying Model)



ภาพประกอบที่ 2-25 สถาปัตยกรรมทั่วไปในการสืบค้นข้อมูล

จากภาพประกอบที่ 2-25 จะแสดงภาพรวมของระบบ เริ่มจากผู้ใช้จะทำการสร้างคำถามและสิ่งที่ต้องการคำตอบผ่านทางรูปแบบการสืบค้นต่อผู้ใช้ หลังจากนั้นระบบจะติดต่อระหว่างส่วนการรวบรวมและสืบค้นข้อมูล และฐานข้อมูลเพื่อประมวลคำตอบ หลังจากนั้นคำตอบจะถูกส่งไปยังรูปแบบการสืบค้น เพื่อแสดงผลของการสืบค้นต่อผู้ใช้

การวัดประสิทธิภาพของการสืบค้น

Precision and Recall (Richard, 2000 และ Cyril and Eric, 2005) คือวิธีการหนึ่งสำหรับการประเมินประสิทธิภาพการสืบค้นเอกสาร ซึ่งสามารถอธิบายได้ ดังนี้

การประเมินผลระบบค้นคืนสารสนเทศเป็นการพิจารณาดูว่า ระบบค้นคืน สารสนเทศจะมีประสิทธิผลมากน้อยเพียงใด ซึ่งประสิทธิภาพของระบบจะวัดได้จากการใช้ทรัพยากรของระบบ เช่น การใช้เนื้อที่หน่วยความจำ, การใช้ CPU Time เป็น ต้น ส่วนประสิทธิผลการวัดด้วยค่า Recall และ Precision ทำได้โดยใช้สูตรต่อไปนี้คือ

$$\text{Recall} = \frac{TruePositive}{TruePositive+FalseNegative}$$

เมื่อ TruePositive หมายถึง จำนวนข้อมูลที่ทำนายถูกว่าเป็นคลาสซึ่งกำลังสนใจอยู่

FaleseNegative หมายถึง จำนวนข้อมูลที่ทำนายผิดมาเป็นคลาสซึ่งไม่ได้สนใจอยู่

หาค่าการแม่นยำของโมเดล (Precision) (เอกสิทธิ์ พัชรวงศ์ศักดา, 2557) โดยพิจารณาแยกทีละคลาส จากสูตร

$$\text{Precision} = \frac{TruePositive}{TruePositive+FalsePositive}$$

เมื่อ TrueNegative หมายถึง จำนวนข้อมูลที่ทำนายถูกว่าเป็นคลาสซึ่งไม่ได้สนใจอยู่

FalsePositive หมายถึง จำนวนข้อมูลที่ทำนายผิดมาเป็นคลาสซึ่งกำลังสนใจอยู่

เราจะทำการประเมินผลระบบเมื่อต้องการเปรียบเทียบความสามารถของ ระบบที่มีอยู่กับระบบอื่น ๆ เมื่อส่วนประกอบบางส่วนของระบบได้เปลี่ยนไป และเมื่อมีการนำเอาส่วนประกอบใหม่ของระบบเข้ามารวมกับระบบที่มีอยู่เดิม

การที่จะประเมินผลระบบนั้น จำเป็นต้องมีการทดสอบระบบ การวัด ความสามารถการทำงานของระบบจะต้องใช้หลักเกณฑ์การวัดที่สามารถแสดงเป็น ปริมาณตัวเลขได้ เราเรียกว่า Objective Measurement ซึ่งจะหาได้โดยการบันทึก, การสังเกตโดยตรงซึ่งจะใช้แบบสอบถาม, เทคนิคในการสัมภาษณ์ หรือใช้เครื่องกลใน การรวบรวมข้อมูล ในกรณีที่ไม่อาจจะหาค่าของพารามิเตอร์ได้ เราก็อาจจะคาดคะเน ด้วยการใช้ Sampling Techniques

ส่วนประกอบของระบบค้นคืนสารสนเทศที่มีอิทธิพลต่อการทำงานของระบบ ได้แก่ ข้อมูลนำเข้า, โครงสร้างของแฟ้มข้อมูล, ภาษาดัชนี, วิธีการสร้างดัชนี, การแสดงรูปแบบของเอกสาร, การวิเคราะห์ข้อคำถาม,วิธีการค้นหาข้อมูลและรูปแบบของผลลัพธ์ที่ได้จากระบบ

ผลการศึกษาที่เกี่ยวข้อง

จากผลการศึกษาค้นคว้า ออนโทโลยี : ทางเลือกของการพัฒนาฐานความรู้ในรูปแบบเชิงเนื้อหา (Ontology: alternatives of knowledgebase development in content base form) ที่ใช้ในการพัฒนาฐานความรู้ที่มีลักษณะเชิงเนื้อหา (Content base) ซึ่งเป็นรูปแบบความรู้ในเชิงประสบการณ์ปัญหาและการแก้ไข แตกต่างจากการวิเคราะห์ความรู้ในรูปแบบของปัญญาประดิษฐ์แบบอื่นๆ ที่มีลักษณะเป็นกฎเกณฑ์ และ/ หรือเป็นข้อมูลเชิงคำนวณออนโทโลยีมีรูปแบบ แนวคิดในการออกแบบและการพัฒนาที่แตกต่างกัน แนะนำความรู้เกี่ยวกับนิยามประโยชน์ รูปแบบ การพัฒนาการปรับปรุง สิ่งที่ต้องคำนึงถึงพร้อมตัวอย่างโดยสังเขป และนำเสนอข้อคิดที่ใช้ในการตัดสินใจเลือกแนวทาง และเครื่องมือมาใช้ในการพัฒนาออนโทโลยีที่เหมาะสมกับความต้องการ เนื่องจากปัจจุบันออนโทโลยียังอยู่ในระหว่าง การพัฒนายังไม่มีการกำหนดมาตรฐานในแต่ละกิจกรรมที่เป็นที่สิ้นสุดและยอมรับโดยรวม ทั้งนี้ เพื่อให้นักวิจัยมีการค้นคว้าหาทฤษฎี แนวทางที่มีประสิทธิภาพสูงสุด

การจัดเก็บและสืบค้นรูปภาพเชิงความหมายบนฐานของ RDF (นพคุณ บุญสิน, 2547) งานวิจัยนี้จะนำเสนอตัวแบบการอธิบายข้อมูลรูปภาพธรรมชาติ โดยตัวแบบจะเป็นโครงสร้างข้อมูลสำหรับอธิบายคุณลักษณะของรูปภาพ ซึ่งประกอบด้วย ประเภทของรูปภาพหมวดของธรรมชาติและหมวดย่อยของธรรมชาติ ตัวแบบดังกล่าวสามารถอธิบายด้วยกราฟบนฐานของ RDF ในงานวิจัยนี้มีการใช้ RDF ในการอธิบายรายละเอียดของรูปภาพ โดยมีขั้นตอนที่สำคัญดังต่อไปนี้

การออกแบบโครงสร้างข้อมูล

เพื่ออธิบายข้อมูลของรูปภาพ โดยแบ่งออกเป็น 4 ระดับเชิงความหมาย ในระดับชั้นที่ 1 คุณสมบัติประเภทของรูปภาพ (Image Type) เป็นคุณสมบัติที่ออกแบบเพื่ออธิบายประเภทของรูปภาพ เช่น ภาพธรรมชาติ ภาพคน และภาพกีฬา เป็นต้น ระดับชั้นที่ 2 คุณสมบัติหมวดของธรรมชาติ (Nature Type) คุณสมบัติที่ออกแบบเพื่ออธิบายหมวดของธรรมชาติที่ปรากฏอยู่ในรูป เช่น ภัยพิบัติ (Disaster) ภูมิประเทศ (Geography) เป็นต้น ระดับชั้นที่ 3 คุณสมบัติหมวดย่อยธรรมชาติ (Nature Category) เป็นคุณสมบัติที่ออกแบบเพื่ออธิบายหมวดย่อยของธรรมชาติในรูปภาพ เช่น หมวดย่อยของภัยพิบัติประกอบด้วย น้ำท่วม ไฟไหม้และแผนดินไหว เป็นต้น ระดับชั้นที่ 4 คุณสมบัติชื่อ (Name) เป็นคุณสมบัติที่ออกแบบเพื่ออธิบายชื่อเฉพาะของหมวดย่อยธรรมชาติในรูป เช่น สึนามิ (Tsunami) อันดามัน (Andaman) เป็นต้น

นิยามออนโทโลยี

คำศัพท์เกี่ยวกับธรรมชาติ ในการจัดเก็บและสืบค้นเชิงความหมายจำเป็นจะต้องนิยามออนโทโลยีคำศัพท์เกี่ยวกับธรรมชาติที่มีความหมายเหมือนกันเพื่อใช้ในการจัดเก็บและสืบค้นรูปภาพ โดยจะใช้ RDF Schema ในการอธิบายข้อมูลของออนโทโลยี

การนำเข้าข้อมูล

เป็นการจัดเตรียมข้อมูลก่อนดำเนินการสืบค้น ซึ่งถือเป็นขั้นตอนที่มีความ สำคัญงานวิจัยนี้จึงได้ออกแบบการนำเข้าข้อมูล โดยมีขั้นตอนดังนี้

1. การสร้าง Lexical Chains จากคำอธิบายรูปภาพ ซึ่งนำเข้าโดยผู้นำเสนอภาพ จะเป็นภาษาธรรมชาติในการอธิบายข้อมูลภาพการสร้าง Lexical Chainsประกอบด้วย 2 ขั้นตอนหลักคือ

1.1 การหาคำสำคัญ จากตัวอย่างคำอธิบายรูปภาพ “While parts of the west coast of Thailand were effect by the 2004 disaster tsunami, most tourist areas had reopened within weeks or even days of the event. Aside from some isolated sports, tourism infrastructure is now intact and business are operating as usual. On badly damaged Koh Phuket.” จะได้คำสำคัญคือ {parts, coast, Thailand, disaster, Tsunami, tourism, areas spots, infrastructure, intact, businesses, Koh, Phuket}

1.2 การสร้าง Lexical Chains จากคำสำคัญดังกล่าวจะนำไปสร้าง Lexical Chains โดยดำเนินการตามลำดับขั้นตอนของการทำ Lexical Chains

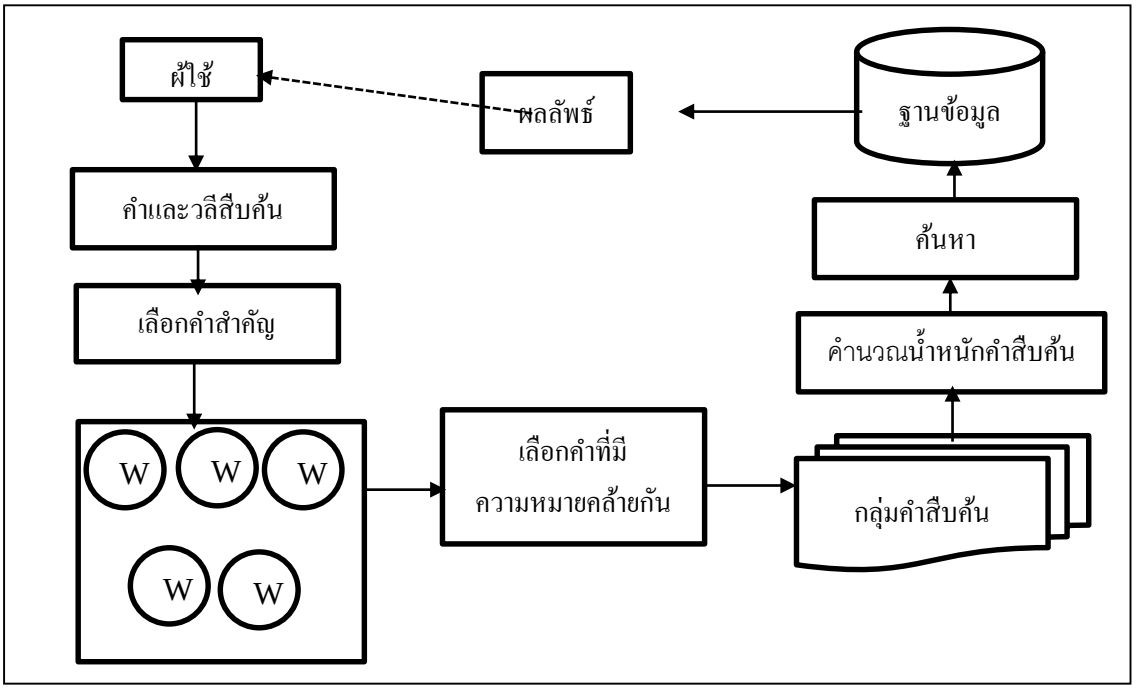
2. สิ่งที่ได้จากขั้นตอนที่ 1 จะนำข้อมูลดังกล่าวไปอธิบายตามโครงสร้างที่ได้ออกแบบไว้ ซึ่งในขั้นตอนนี้มีการใช้ Ontology จากคำศัพท์ธรรมชาติเข้ามาร่วมด้วยคล้ายกับการสร้าง Lexical Chains

3. นำข้อมูลไปอธิบายด้วยภาษา RDF

4. จัดเก็บข้อมูลในฐานข้อมูลในรูปแบบของเอกสาร RDF ที่เชื่อมโยงกับรูปภาพ

การออกแบบการสืบค้นรูปภาพ

การสืบค้นข้อมูลโดยการใช้คำสำคัญ (Keyword) และคำสืบค้นด้วยภาษาธรรมชาติ (Natural Language) ล้วนแต่มีประสิทธิภาพต่ำ เนื่องมาจากข้อมูลที่สืบค้นมีโครงสร้างที่ไม่แน่นอน จึงทำให้การสืบค้นข้อมูลได้ผลลัพธ์ไม่ตรงกับความต้องการ การออกแบบขั้นตอนการสืบค้นข้อมูลเป็นขั้นตอนสำคัญในการสืบค้นข้อมูล โดยการออกแบบขั้นตอนจะดำเนินตามโครงสร้างข้อมูลที่ได้กำหนดขึ้น เริ่มจากนำคำหรือวลีสืบค้นมาเลือกสำคัญ ต่อจากนั้นจะนำคำสำคัญดังกล่าวมาค้นหาคำที่มีความหมายคล้ายกัน (Synonym) จาก Ontology คำศัพท์ธรรมชาติ หลังจากนั้นจะทำการคำนวณน้ำหนักของกลุ่มคำสืบค้น และนำกลุ่มคำสำคัญไปค้นหาในฐานข้อมูล



ภาพประกอบที่ 2-26 แสดงขั้นตอนการสืบค้นข้อมูล

1) การเลือกคำสำคัญจากคำหรือวลีสืบค้น จากคำหรือวลีสืบค้นที่ผู้ใช้ ส่งเข้าสู่ระบบเพื่อสืบค้นข้อมูล คำหรือวลีดังกล่าวเป็นภาษาธรรมชาติ (Natural Language) ซึ่งการค้นหาด้วยรูปแบบดังกล่าวทำให้ผลลัพธ์จากการสืบค้นไม่ตรงกับความต้องการ ดังนั้นจึงจำเป็นต้องมีการจัดรูปแบบโดยการการพิจารณาเลือกเฉพาะคำสำคัญที่สื่อความหมายได้มาเป็นคำสำคัญในการสืบค้นโดยอาศัยฐานข้อมูล Ontology คำศัพท์ธรรมชาติ ตัวอย่าง วลีสืบค้น “Picture of disaster about tidalwave tsunami in koh phuket.” จากวลีดังกล่าวจะเลือกคำสำคัญได้ดังนี้ {disaster, tidalwave tsunami koh phuket} นำคำสำคัญดังกล่าวเลือกคำศัพท์ที่มีความหมายเหมือนกัน (Synonym) ความสำคัญของคำที่มีความหมายเหมือนกัน เนื่องจากการสืบค้นด้วยคำสำคัญ เป็นการเปรียบเทียบคำสำคัญกับข้อมูลในฐานข้อมูล ถ้าการเปรียบเทียบไม่พบข้อมูลแสดงว่าไม่มีข้อมูลที่ตรงกับความต้องการของผู้ใช้ สำหรับการเลือกคำที่มีความหมายเหมือนกันเพื่อเป็นกลุ่มคำในการสืบค้นข้อมูล มีจุดประสงค์คือทำการสืบค้นข้อมูลด้วยคำที่มีความหมายเหมือนกันหรือใกล้เคียงกันจะทำให้การสืบค้นมีประสิทธิภาพและเชิงความหมายด้วย ภาพประกอบที่ 3.18 เป็นการนิยามคำศัพท์ที่มีความความหมายเหมือนกันโดยภาษา RDF Schema

2) การคำนวณน้ำหนักของคำสืบค้น จากกลุ่มคำสืบค้นที่ได้จากขั้นตอนที่ 3.6.2 ในขั้นตอนนี้เป็นการคำนวณน้ำหนักของคำสืบค้นเพื่อเป็นคะแนนในเปรียบเทียบและสืบค้นรูปภาพ โดยการคำนวณน้ำหนักของคำสืบค้น จะอาศัยฐานข้อมูล Ontology ของคำศัพท์ธรรมชาติและน้ำหนักของโครงสร้างที่ได้นิยามขึ้นในขั้นตอนที่ 3.2 ซึ่งมีการให้น้ำหนักตามลำดับดังนี้ ระดับที่ 1 2 3 และ 4 มีน้ำหนักคือ 1 2 4 และ 8 ภาพประกอบที่ 2-19 เป็นตัวอย่างการคำนวณน้ำหนักของคำสืบค้น โดยในคำสืบค้นประกอบด้วย คำในระดับที่ 3 (1 คำ) และคำระดับที่ 4 (2 คำ) ผลรวมของน้ำหนักเท่ากับ 20

3) การสืบค้นข้อมูลจากฐานข้อมูล เมื่อทำการจัดรูปแบบข้อมูลนำเข้าและคำสืบค้นเรียบร้อยแล้ว ขั้นตอนต่อมาคือการออกแบบการสืบค้นข้อมูลจะดำเนินการตามโครงสร้างข้อมูลเพื่อให้ได้ผลลัพธ์ตรงกับความต้องการของผู้ใช้

```
MF(Query[Q1,Q2,…,Qn], Picture[P1,P2,…,Pm],QW)
{
i=0, j=0;
PW = 0;
For i from 0 to n
For j from 0 to m
Switch (ontology_type (Qi)) ;แยกประเภทเปรียบเทียบตามคุณสมบัติ
{
Case ImageType ; บวกค่าเพมิ่ ตามน้ำหนักของแต่ละคุณสมบัติ
If (Qi = Pj) {PW=PW+1 }
Case NatureType
If (Qi = Pj) {PW=PW+2 ; }
Case NatureCategory
If (Qi = Pj) {PW=PW+4 ; }
Case Name
If (Qi = Pj) {PW=PW+8; }
}
If (PW >= ½(QW)) ; เปรียบเทียบน้ำหนักของรูปภาพ กับ น้ำหนัก
ของคำสืบค้น
{ Return “Picture [name]” ;ชื่อของรูป หรือที่อยู่ของรูป}
}
```

ภาพประกอบที่ 2-27 ลำดับขั้นตอนการการสืบค้นข้อมูล

ลำดับขั้นตอนการสืบค้นข้อมูลตามภาพประกอบที่ 2-27 มีลำดับขั้นตอนการสืบค้นข้อมูลดังนี้

3.1 นำข้อมูลเข้าสู่ขั้นตอนการสืบค้น (Matching Function = MF) ข้อมูลประกอบ ด้วยกลุ่มคำสืบค้น (Query [Q1,Q2,…,Qn]) โดย Q1,Q2,…,Qn คือกลุ่มคำสืบค้น ข้อมูลของรูปภาพ (Picture [P1,P2,…,Pm]) โดย P1,P2,…,Pm คือคุณสมบัติของรูปภาพและน้ำหนักของคำสืบค้น (Query Weight = QW)

3.2 แยกประเภทของกลุ่มคำสืบค้นโดยใช้ Ontology คำศัพท์ธรรมชาติ

3.3 คำนวณน้ำหนักของรูปภาพ (Picture Weight =PW) ตามคุณสมบัติที่ค้นพบ โดยแต่ละคุณสมบัติจะมีน้ำหนักดังนี้ ระดับที่ 1 2 3 และ 4 มีน้ำหนักคือ 1 2 4 และ 8

3.4 นำค่าน้ำหนักของรูปภาพที่ได้ มาทำการเปรียบเทียบกับน้ำหนักของกลุ่มคำสืบค้น ถ้าน้ำหนักของภาพที่ค้นพบ มากกว่าร้อยละ 50 ของน้ำหนักกลุ่มคำสืบค้น แสดงภาพภาพดังกล่าว มีความถูกต้องและตรงกับความต้องการ การเปรียบเทียบคะแนนของน้ำหนักของการสืบค้นและน้ำหนักของรูปภาพ ถ้าคะแนนของรูปภาพมีน้ำหนักน้อยกว่าร้อยละ 50 ของคะแนนน้ำหนักของคำสืบค้นจะทำให้ได้ผลลัพธ์ไม่ตรงกับความต้องการ ถ้าคะแนนรวมของน้ำหนักมากกว่าร้อยละ 50 โดยเฉลี่ยจะได้ผลลัพธ์ที่ตรงกับความต้องการ

บทที่ 3

วิธีการดำเนินการศึกษา

การดำเนินการศึกษา

ในการดำเนินการศึกษามีวัตถุประสงค์เพื่อศึกษาการนำทฤษฎีออนโทโลยีมาใช้ในการจำแนกลักษณะของภาพ Microstock เพื่อช่วยให้การทำงานมีประสิทธิภาพในการวิเคราะห์ข้อมูลถูกต้องและง่ายต่อการใช้งาน โดยการนำเอาเครื่องมือในการสนับสนุนการวิเคราะห์มาใช้ลดความซ้ำซ้อนของข้อมูล และใช้ในการศึกษาเพื่อสร้างระบบธุรกิจ Microstock ที่มีการตอบสนองความต้องการของผู้ใช้ การแยกหมวดหมู่ของภาพอย่างชัดเจนเพื่อความสะดวกในการค้นหาและได้ภาพที่ตรงกับความต้องการที่สุด เหมาะสมกับความต้องการของผู้ใช้แต่ละบุคคลต่อไป โดยออนโทโลยีที่ใช้มีความสำคัญต่อการจัดสร้าง การใช้งานฐานความรู้เชิงเนื้อหาขององค์กรได้อย่างมีประสิทธิภาพ การจัดสร้างฐานความรู้ต้องคำนึงถึงความถูกต้องของข้อมูล ในการจัดเก็บ และจากการสืบค้นหาคำตอบที่ถูกต้อง การจัดเตรียมบุคลากรเพื่อทำหน้าที่จัดสร้าง ดูแล บำรุงรักษา การเลือกเครื่องมือ (Computer language-Tool) เป็นสิ่งที่พึงพิจารณาความสามารถในการอนุมานความรู้ได้ลึกซึ้งขึ้น และสามารถทำงานได้ต่าง Platform หรือ Interoperability จะทำให้สามารถเชื่อมโยง Collaborative กับออนโทโลยีอื่น ๆ ได้ดียิ่งขึ้น การทำงานในยุคปัจจุบันจะมีการปฏิบัติงานบนรูปแบบการประมวลผลแบบกระจาย (Distributed processing) บนเครือข่ายอินเทอร์เน็ต นั่นคือการพิจารณาการรักษาความปลอดภัยของฐานความรู้เป็นสิ่งที่จำเป็นสำคัญ เพื่อให้ปลอดภัยจากการถูกบุกรุกฐานความรู้หรือคลังสมองของระบบ

เครื่องมือและอุปกรณ์ที่ใช้

(1) ภาษา SPARQL สำหรับการเขียนคิวรี (Query) ข้อมูลจากฐานความรู้ออนโทโลยี (นฤพนธ์ พนาวงศ์, 2553)

(2) โปรแกรม Protégé’ 5.0.1 สำหรับสร้างฐานความรู้ออนโทโลยีตามกฎ OWL-DL (Protégé Project, 2016)

(3) ใช้แบบสอบถามหาค่าความตรงตามเนื้อหา

วิธีดำเนินการ

การสร้างพจนานุกรมคำศัพท์

ผู้วิจัยได้สร้างพจนานุกรมคำศัพท์ที่มีความหมายเกี่ยวข้องกับรูปภาพในลักษณะต่างๆ เพื่อใช้เป็นฐานข้อมูลคำศัพท์ที่เกี่ยวข้องกับรูปภาพ สำหรับนำไปใช้ในการเปรียบเทียบกับคำค้นที่ผู้ใช้ป้อนเข้ามาในระบบ และเพื่อช่วยให้ระบบเข้าใจว่าจากคำค้นนั้นๆ ควรจะใช้ลักษณะการค้นหาแบบใด เพื่อให้ผลลัพธ์ที่ได้จากการเปรียบเทียบมีประสิทธิภาพ และช่วยให้ใช้เวลาในค้นหาเร็วขึ้น

ตารางที่ 3.1 ตารางตัวอย่างพจนานุกรมคำศัพท์ จากการจำแนกประเภทภาพและกำหนดคำศัพท์ (Vocabulary)

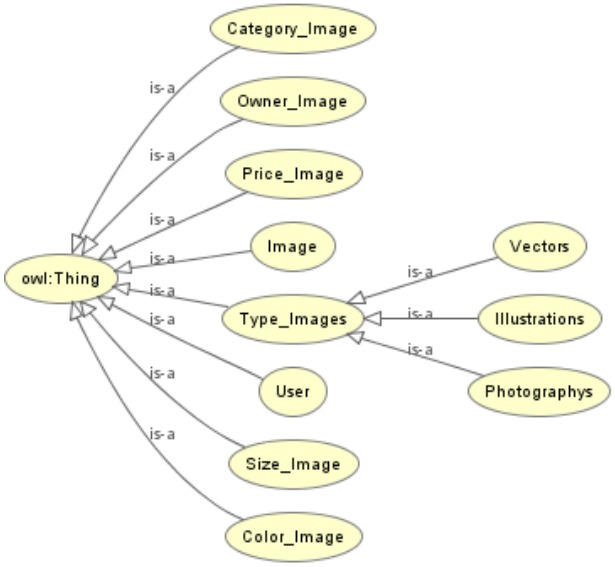
ประเภท	หมวดหมู่	สี	ขนาด	ราคา
Illustrations	Animals	Any	640x480	10USD
Photographys	BeautifulWomen	B/W	800x600	FREE
Vectors	Buildings	Full	1024x768	
	ClipArt	Green	1600x1200	
	Emotion	Orange	2272x1704	
	Flower	Pink	Exactly	
	Food	Red	Icon	
	Sport	Transparent	Lange	
	Painting	Yellow	Medium	

คำศัพท์ในพจนานุกรมคำศัพท์ แบ่งออกเป็น 5 ประเภท คือ ประเภทของภาพ หมวดหมู่ของภาพ โทนสีของภาพ ราคา ขนาด เจ้าของภาพ คำสำคัญ โดยแยกคำศัพท์ตามความสัมพันธ์จากฐานข้อมูลออนโทโลยีรูปภาพ ดังตัวอย่างในตารางที่ 3.1

การออกแบบคอนเซพของฐานความรู้ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

คอนเซพของฐานความรู้ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์นั้นประกอบด้วยคลาสหลัก ดังต่อไปนี้

- คลาสประเภทของรูปภาพ (Category_Image)
- คลาสหมวดหมู่ของรูปภาพ (Type_Image)
- คลาสโทนสีของรูปภาพ (Color_Image)
- คลาสราคาของรูปภาพ (Price_Image)
- คลาสขนาดของรูปภาพ (Size_Image)
- คลาสเจ้าของรูปภาพ (Owner_Image)



ภาพประกอบที่ 3-1 แสดงการออกแบบออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

ภาพประกอบที่ 3-1 แสดงการออกแบบออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์โดยที่งานวิจัยนี้ได้ออกแบบออนโทโลยีสำหรับการค้นหารูปภาพที่สร้างด้วยโปรแกรม Protégé 5.0.1 โดยใช้เครื่องมือ OWLViz

หลังจากที่ได้ออกแบบออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์ ผู้วิจัยได้นำข้อมูลของภาพมาบรรยายให้อยู่ในรูปแบบของ RDF (Resource Description Framework) โดยใช้ OWL (Ontology Web Langrage) ที่ได้มีการบรรยายข้อมูลของรูปภาพตามโครงสร้างของออนโทโลยีที่ได้ออกแบบไว้ เพื่อรวบรวมไว้เป็นฐานความรู้สำหรับนำไปใช้ในการสืบค้นต่อไป

ในการนิยามโครงสร้างการอธิบายข้อมูลรูปภาพนั้น เป็นการอธิบายเพื่อให้มีลักษณะของภาษาเว็บเชิงความหมาย ซึ่งประกอบด้วยคุณสมบัติต่างๆ ทั้งคุณสมบัติของวัตถุและคุณสมบัติของชนิดข้อมูล เช่น

- หมวดหมู่ประเภทของรูปภาพ และชื่อของรูปภาพ
- หมวดหมู่ลักษณะของรูปภาพ
- หมวดหมู่ของสีของรูปภาพ
- หมวดหมู่ของขนาดของรูปภาพ ฯลฯ

โดยที่ข้อมูลเหล่านี้ถูกแสดงด้วยเซตของทริปเปิล <subject, predicate, object>

เมื่อ Subject หมายถึง ทรัพยากร (Resource)

Predicate หมายถึง คุณสมบัติ (Property)

Object หมายถึง ค่าของคุณลักษณะ (Property value)

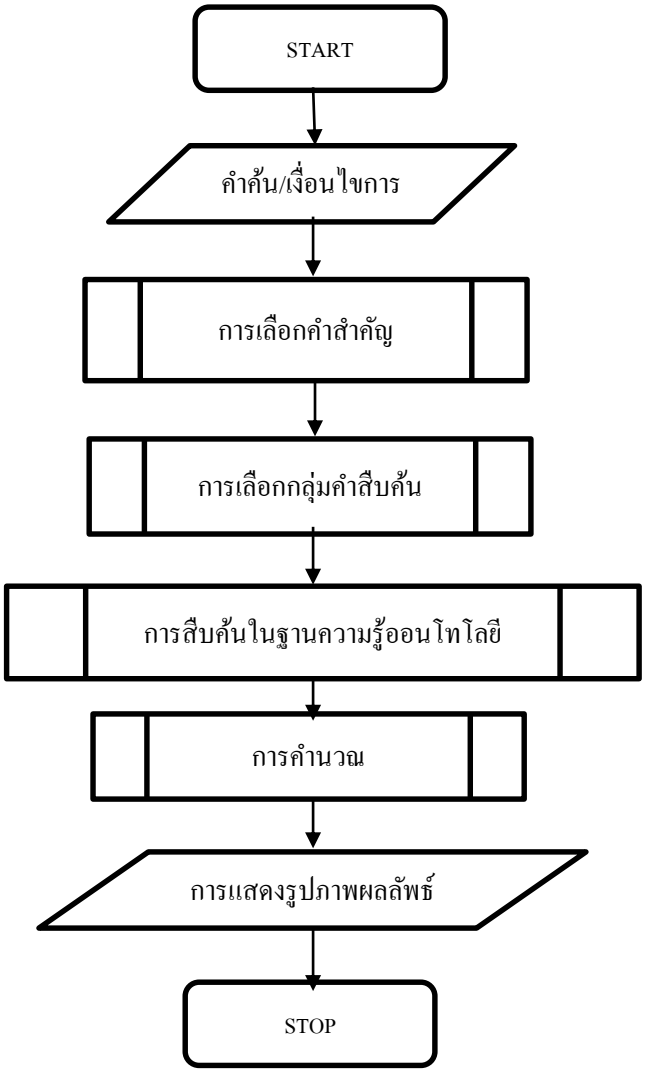
```
<!-- http://www.semanticweb.org/ontologies/Imagesearch.owl#Cat -->
<owl:Thing rdf:about="&Imagecategory;Cat">
< Imagecategory:Featured
>lovely sleep cat
</ Imagecategory:Featured>
< Imagecategory:image>Animal\Cat</ Imagecategory:image>
<Name>Sleep cat</Name>
<Imagetourism:Caused>brown cat
</Imagetourism:Caused>
<Imagetourism:hasRelation rdf:resource="&Imagetourism;Cat "/>
<Imagetourism:hasMaterial rdf:resource="&Imagetourism;Sleepcat"/>
<Imagetourism:hasOccur rdf:resource="&Imagetourism;Animal"/>
<hasContact rdf:resource="#Photography "/>
<hasClassification rdf:resource="#Animal "/>
```

ภาพประกอบที่ 3-2 ตัวอย่างการอธิบายข้อมูลรูปภาพด้วยภาษา OWL

ภาพประกอบที่ 3-2 แสดงตัวอย่างการอธิบายข้อมูลรูปภาพด้วยภาษา OWL ซึ่งประกอบด้วยการนิยามหมวดหมู่ของรูปภาพ (Attraction) ซึ่งมีซับคลาส (SubClass) คือ ประเภทของภาพ (Photography) และมีซับคลาส (SubClass) คือ ประเภทสัตว์ (Animal) ซึ่งเป็นประเภทหมวดหมู่ของรูปภาพที่มีความสัมพันธ์กับภาพ “Sleep cat” และได้รับการอธิบายคำค้นจาก Data Properties ของ Name และ Feature

การสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยี

ระบบการสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยี ประกอบไปด้วยขั้นตอนการทำงาน 4 ขั้นตอน คือ การเลือกคำสำคัญ การเลือกกลุ่มคำสืบค้น การค้นหาในฐานความรู้ออนโทโลยีและการคำนวณค่าน้ำหนักดังภาพประกอบที่ 3-3



ภาพประกอบที่ 3-3 ขั้นตอนการทำงานของระบบ

การเลือกคำสำคัญ

ผู้วิจัยได้ออกแบบและพัฒนาระบบการสืบค้นรูปภาพเชิงความหมาย ที่กำหนดให้ผู้ใช้งานระบบสามารถสืบค้นข้อมูลที่ต้องการ ด้วยการระบุคำค้นที่ต้องการ หรือระบุเงื่อนไขที่ต้องการสืบค้นข้อมูลที่ต้องการ

กรณีที่ผู้ใช้ต้องการสืบค้นข้อมูลด้วย คำ ค้น ระบบนำ คำ ค้นที่ผู้ใช้ป้อนเข้าระบบเปรียบเทียบกับคำศัพท์ในพจนานุกรมคำศัพท์ที่มีความหมายเกี่ยวข้องกับรูปภาพ ผู้วิจัยได้แบ่งคำศัพท์ในพจนานุกรมคำศัพท์ออกเป็น 6 ประเภท คือ

1) ประเภทของรูปภาพ

2) ลักษณะของรูปภาพ

3) ราคาของรูปภาพ

4) สีของรูปภาพ

5) ขนาดของรูปภาพ

6) คำอธิบายของรูปภาพ

ถ้าคำค้นที่ผู้ใช้ป้อนตรงกันกับคำศัพท์ในพจนานุกรมคำศัพท์ประเภทที่ 1-5 คือ ประเภทของรูปภาพ ลักษณะของรูปภาพ ราคาของรูปภาพ สีของรูป และขนาดของรูปภาพ จะทำให้ระบบได้ลักษณะการสืบค้นของคำค้นนั้น ๆ ว่าจะเป็นการสืบค้นจากลักษณะของรูปภาพ หรือ การสืบค้นจากประเภทของรูปภาพ หรือการสืบค้นจากราคาของรูปภาพ หรือ การสืบค้นจากสีของรูปภาพ หรือ การสืบค้นจากขนาดของรูปภาพ

ถ้าคำค้นที่ผู้ใช้ป้อนตรงกันกับคำศัพท์ในพจนานุกรมคำศัพท์ประเภทที่ 6 คือ คำอธิบายของรูปภาพ ระบบจะนำคำค้นนั้น มาผ่านกระบวนการตัดคำ โดยใช้โปรแกรมแบ่งคำภาษาไทย SWATH และนำคำที่ตัดได้นี้ ไปเปรียบเทียบกับพจนานุกรมคำศัพท์ ประเภทที่ 1 คือ ประเภทของรูปภาพ ประเภทที่ 4 คือ สีของรูปภาพ และประเภทที่ 6 คือ คำอธิบายของรูปภาพ

ถ้าคำค้นที่ผู้ใช้ป้อนไม่ตรงกันกับคำศัพท์ในพจนานุกรมคำศัพท์ประเภท 1-6 ระบบจะให้ผู้ใช้ป้อนคำค้นเข้าระบบใหม่ หรือให้ผู้ใช้เลือกเงื่อนไขเพิ่มเติมจากระบบ เพื่อให้ระบบทำการสืบค้นในเงื่อนไขที่ตรงกับความต้องการของผู้ใช้มากยิ่งขึ้น

กรณีที่ผู้ใช้ต้องการสืบค้นข้อมูลด้วยการระบุ “เงื่อนไขการสืบค้น” ผู้ใช้ต้องเลือกกลุ่มของเงื่อนไขการสืบค้น โดยที่ระบบได้แบ่งเงื่อนไขการสืบค้นเป็น 2 กลุ่ม คือ

- การสืบค้นตามลักษณะและประเภทของรูปภาพ
- การสืบค้นตามลักษณะเฉพาะของรูปภาพเช่น สี ขนาด เจ้าของรูปภาพ

การเลือกกลุ่มคำสืบค้น

ระบบนำ คำสำคัญที่ผู้ใช้ป้อนเข้ามาในระบบและผ่านการตรวจสอบแล้วว่าอยู่ในพจนานุกรมคำศัพท์ สืบค้นในฐานความรู้ออนโทโลยี เพื่อให้ได้กลุ่มคำสืบค้น ซึ่งเป็นคำที่มีความหมายเหมือนกัน (Synonym) เพื่อให้ระบบนำไปใช้ในการสืบค้นข้อมูลในขั้นตอนต่อไป ซึ่งจะช่วยให้การสืบค้นข้อมูลมีประสิทธิภาพและเชิงความหมายมากยิ่งขึ้น

คำค้นที่ 1 : Grafiled

ตารางที่ 3-2 ตัวอย่างผลลัพธ์ของการเลือกกลุ่มคำสืบค้น

No.	? synonym	?type
1.	Resource: http://www.semanticweb.org/ontologies/ Imagesearch.owl#Garfiled_Cat	Resource: http://www.semanticweb.org/ontologies/ Imagesearch.owl#Cat
2.	Resource: http://www.semanticweb.org/ontologies/tourismm.owl#Sleepy_Cat	Resource: http://www.semanticweb.org/ontologies/ Imagesearch.owl#Cat

ตารางที่ 3-2 เป็นตัวอย่างของการเลือกกลุ่มคำสืบค้น ที่ผู้ใช้ป้อนคำค้น “Grafiled” จากนั้นระบบนำคำค้นไปสืบค้นในส่วนของเอกสาร OWL ได้ผลลัพธ์เป็นคำที่มีความหมายเหมือนกัน คือ คำว่า Sleepy cat และแสดงผลลัพธ์ออกมาให้ทราบว่ากลุ่มคำเหล่านี้มี Resource เป็น http://www.semanticweb.org/ontologies/Imagesearch.owl#Cat

จากนั้นระบบนำกลุ่มคำสืบค้นที่ได้ไปสืบค้นข้อมูลในเอกสาร OWL โดยทำการเปรียบเทียบกับข้อมูลในโหนดที่เป็น Literal ภายใน Resource ที่ระบุไว้ในกลุ่มคำสืบค้นนั้น ๆ เท่านั้น เพื่อเป็นการแก้ปัญหาของการสืบค้นข้อมูลที่ได้ผลลัพธ์จำนวนมาก แต่มีบางผลลัพธ์ไม่ตรงตามความต้องการของผู้ใช้

ตารางที่ 3.3 ตัวอย่างผลลัพธ์ของการสืบค้นข้อมูลรูปภาพจากเอกสาร OWL

No.	? resource	?title
คำค้นที่ 1 : Find Grafiled		
1.	Resource: http://www.semanticweb.org/ontologies/ Imagesearch.owl#Garfiled_Cat	Literal : Orange cat
คำค้นที่ 2 : Find Sleepy cat		
1.	Resource: http://www.semanticweb.org/ontologi es/tourismm.owl#Sleepy_Cat	Literal : Lazy cat

ตารางที่ 3-3 เป็นตัวอย่างผลลัพธ์ของการสืบค้น เมื่อคำค้นที่ผู้ใช้ป้อนเข้ามาคือ “Gar filed” และกลุ่มคำสืบค้นที่ได้จากขั้นตอนการเลือกลุ่มคำสืบค้น ประกอบด้วย 2 คำคือ Grafiled และ Sleepy cat หลังจากนั้นระบบสืบค้นข้อมูลจะนำคำสืบค้นทั้ง 2 คำนี้ไปทำการสืบค้นข้อมูลในโหนดที่มีค่าเป็น Literal ภายใต้ Resource: http://www.semanticweb.org/ontologies/Imagesearch.owl#Cat และได้ผลลัพธ์ออกมาดังนี้

คำค้นที่ 1 “Grafiled” ได้ผลลัพธ์ของโหนดที่มีค่าเป็น Literal ที่บรรจุค่าที่ตรงกับคำสืบค้น คือ Literal : Orange cat ซึ่งเป็นค่าของคุณสมบัติ Title โดยมี Resource เป็น http://www.semanticweb.org/ontologies/Imagesearch.owl#Garfiled_Cat

คำค้นที่ 2 “Sleepy cat” ได้ผลลัพธ์ของโหนดที่มีค่าเป็น Literal ที่บรรจุค่าที่ตรงกับค่าคำสืบค้น คือ

Literal : Lazy cat ซึ่งเป็นค่าของคุณสมบัติ Title โดยมี Resource เป็น http://www.semanticweb.org/ontologies/tourismm.owl#Sleepy_Cat

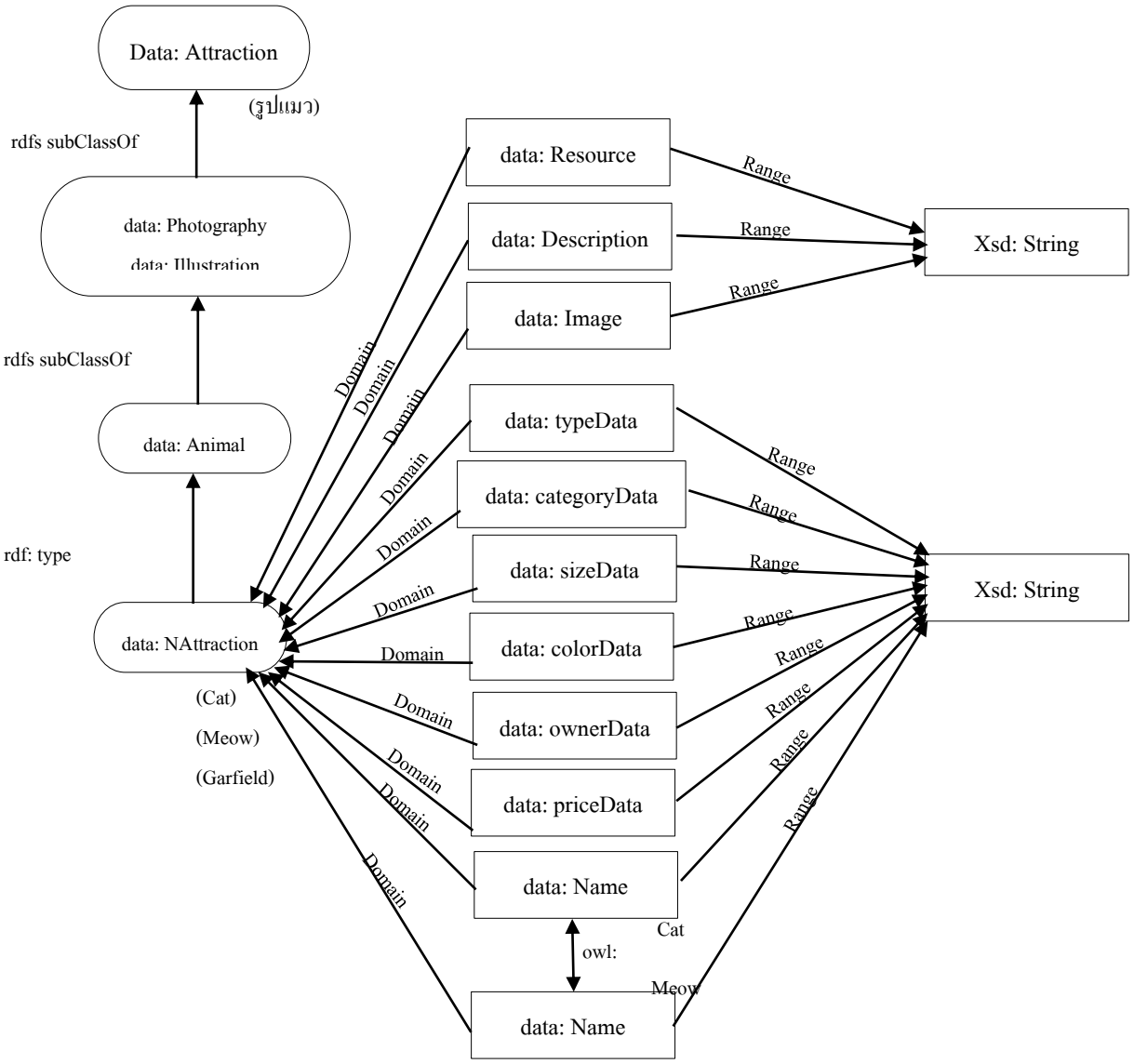
การสืบค้นในฐานความรู้ออนโทโลยี

จากข้อมูลที่มีความสัมพันธ์กันตามโครงสร้างการอธิบายข้อมูล OWL ซึ่งเป็นการอธิบายว่าข้อมูลมีคุณสมบัติ และค่าคุณสมบัติสัมพันธ์ตรงกันกับสิ่งใดบนเงื่อนไขที่มีความสัมพันธ์กับคำค้นที่ผู้ใช้ป้อนเข้ามาในระบบ ด้วยการสืบค้นด้วยภาษา SPARQL

ตารางที่ 3-4 ตัวอย่างข้อมูลที่มีความสัมพันธ์กันจากโครงสร้างการอธิบายข้อมูล OWL

Subject	**Predicate**	**Object**
Garfield	hasContact	Animal
Cat	rdfs:subClassOf	Animal
Garfieldcat.jpg	rdf:type	Cat

ตัวอย่างการสืบค้นข้อมูลรูปภาพเชิงความหมาย



ภาพประกอบที่ 3-4 ตัวอย่างโครงสร้างออนโทโลยีคลาสของรูปภาพ

จากภาพประกอบที่ 3-4 ตัวอย่างการสืบค้นข้อมูลรูปภาพเชิงความหมาย มีขั้นตอนดังนี้

1) กระบวนการทำงาน

เมื่อผู้ใช้ทำการสืบค้นข้อมูลรูปภาพ ด้วยการระบุคำค้น “Garfield” ระบบจะนำคำค้นที่ผ่านการตรวจสอบว่าคำค้นอยู่ในพจนานุกรมคำศัพท์ไปดำเนินการสืบค้นในฐานความรู้ออนโทโลยีโดยมีขั้นตอนการทำงานดังนี้

1. ระบบทำการสืบค้นโดยคำค้นจาก Data Properties ชื่อ Name ซึ่งเมื่อใส่คำค้นเข้าไป เช่น คำ Garfield คำค้นนั้นจะตรวจสอบว่ามี Individuals ใดๆ ที่มีคุณสมบัติเป็น owl:sameAs ของ Individuals ที่เป็นคำค้นนั้นๆ จากคุณสมบัติของ owl:sameAs จะทำให้กระบวนการค้นหาที่มีความหมายเหมือนกันแต่เขียนต่างกัน เช่นคำว่า Cat Meow ต่างก็มีความหมายเดียวกับ คำว่า Garfiled

2. ใช้ Data Properties ชื่อ Name ที่อยู่ในหมวดหมู่อยู่ในคลาสรูปภาพ ที่มีค่า Individuals คือ Garfield หรือ Cat ซึ่งหมายถึงในกรณีที่ผู้ใช้ต้องการค้นหารูปสถานที่ไม่ว่าจะเป็น “Garfield” หรือ “Cat” นั่นหมายถึง 2 ชื่อนี้คือสิ่งเดียวกัน

3. จากนั้นระบบจะทำการค้นหาไปยังโหนดต่าง ๆ ที่อยู่ภายในคลาสเดียวกันกับหมวดหมู่ที่ได้ออกแบบไว้เพื่ออธิบายคุณสมบัติของรูปภาพนั้น ๆ เช่น ประเภทของภาพ (TypeData) รูปภาพ (Image) คำอธิบาย (Description) ลักษณะ (categoryData) เป็นต้น เพื่อแสดงผลลัพธ์จากการค้นหาที่ตรงกับคำค้น

2) ตัวอย่างการแทนค่าข้อมูล

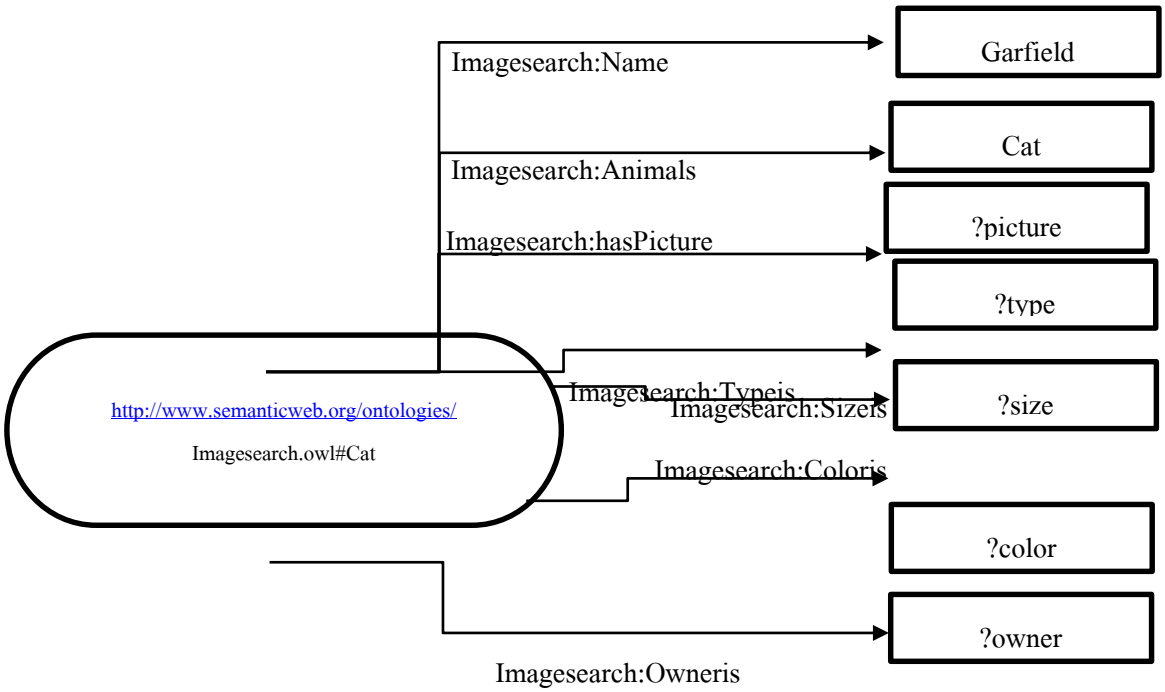
เป็นการแทนค่าข้อมูลเพื่ออธิบายข้อมูลในโครงสร้างการอธิบายข้อมูล OWL เพื่ออธิบายว่าข้อมูลมีคุณสมบัติ และค่าคุณสมบัติสัมพันธ์กับสิ่งใด ซึ่งข้อมูลจะถูกแทนเป็น Resource แล้วแทนค่า Resource ด้วย URI แล้วอธิบาย Resource ด้วย Property และค่าของ Property โดยใช้ Metadata

การแทนข้อมูลโครงสร้างการอธิบายข้อมูล OWL จะใช้วิธีในการบรรยายรายละเอียดของข้อมูลที่เรียกว่า Triple โดยประกอบด้วย 3 ส่วนคือ

Subject คือ ส่วนทรัพยากรหรือส่วนสิ่งที่สนใจ ซึ่งเรียกว่า Resource

Predicate คือ ส่วนที่บ่งบอกคุณสมบัติของ Resource

Object คือ ส่วนที่เป็นค่าของคุณสมบัติของ Resource



ภาพประกอบที่ 3-5 ตัวอย่างการบรรยายรายละเอียดข้อมูลของ OWL

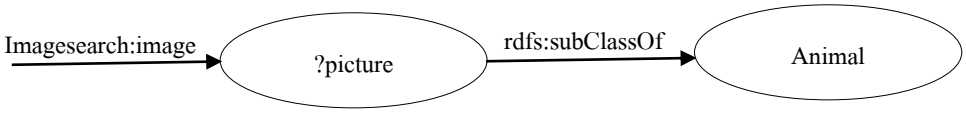
ภาพประกอบที่ 3-5 แสดงตัวอย่างการบรรยายรายละเอียดข้อมูลของ OWL ที่เรียกว่า Triple เมื่อนำตัวอย่างข้อมูลรูปภาพมาทำการแทนค่าข้อมูลในโครงสร้างการอธิบายข้อมูล OWL และนำมาแสดงในรูปแบบกราฟ เพื่อแสดงตัวอย่างการอธิบายข้อมูลรูปภาพที่มีความสัมพันธ์กันในหมวดประเภทของรูปภาพ

ตัวอย่างการค้นหาข้อมูลรูปภาพด้วย SPARQL

ภาพประกอบที่ 3-6 แสดงตัวอย่างการค้นหาข้อมูลรูปภาพจากคลาสภาพถ่าย ซับคลาสสัตว์ ประกอบด้วย ข้อมูลรูปภาพสัตว์ต่างๆ ทำให้ได้ผลลัพธ์จากการสืบค้น คือ Dog Scottydoo Cat Meow Garfield เป็นต้น

```
SELECT ?x
        WHERE { ?x rdfs:subClassOf: Animal}
                        Dog
                        Scoobydoo
                        Cat
                        Meow
                        Garfield
SELECT? Picture
        WHERE { ?x rdfs:subClassOf seach:Animal;
                        ?x Imagesearch:image ?picture.
```

ภาพประกอบที่ 3-6 ตัวอย่างการค้นหาข้อมูลรูปภาพจากซับคลาสหมวดลักษณะของรูปภาพ



ภาพประกอบที่ 3-7 อธิบายการค้นหาข้อมูลรูปภาพจากซับคลาสหมวดสัตว์

กำหนดซับคลาส rdfs:subClassOf เป็นการกำหนดคลาสใหม่ ซึ่งเกิดการถ่ายทอดคุณสมบัติของคลาส กำหนดให้ค้นหารูปภาพที่เป็นซับคลาสของคลาส Attraction (รูปภาพ) คือ Animal แสดงผลรูปภาพจากเงื่อนไขคุณสมบัติ Data Properties : Imagesearch:image ผ่านตัวแปร ?picture ซึ่งผลลัพธ์ที่ได้จะเป็นรูปภาพ Dog Scoobydoo Cat Meow Garfield เป็นต้น

ตารางที่ 3-5 ตัวอย่างผลลัพธ์ที่ได้จากการคิวรีข้อมูลด้วยภาษา SPARQL

No.	?picture
1.	
2.	
3.	
4.	
5.	

การคำนวณน้ำหนัก

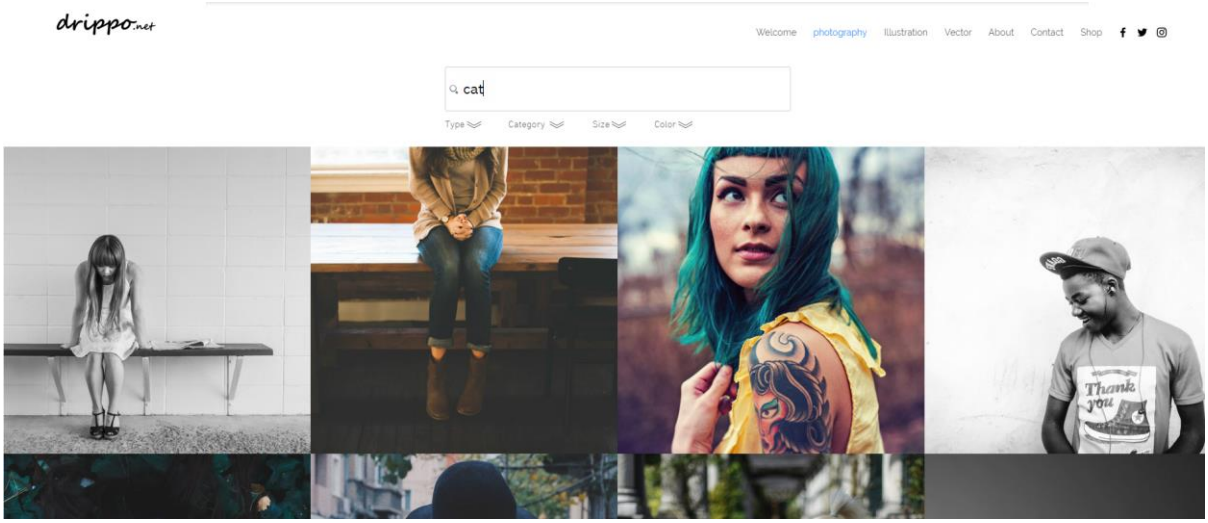
ขั้นตอนนี้ระบบทำการคำนวณน้ำหนักของผลลัพธ์ที่ได้จากการสืบค้นข้อมูลรูปภาพ ซึ่งอาจจะเป็นการสืบค้นโดยใช้คำค้น หรือสืบค้นโดยใช้เงื่อนไขการสืบค้น เพื่อให้ระบบนำค่าน้ำหนักที่คำนวณได้นี้ไปใช้ในการจัดลำดับการแสดงผลลัพธ์ในขั้นตอนต่อไป ในการคำนวณน้ำหนักนั้น ระบบได้แบ่งรูปแบบการคำนวณน้ำหนัก ตามคลาสของฐานข้อมูลออนโทโลยี ที่ประกอบไปด้วย

- คลาสประเภทของรูปภาพ
- คลาสหมวดหมู่ของรูปภาพ
- คลาสโทนสีของรูปภาพ
- คลาสราคาของรูปภาพ
- คลาสขนาดของรูปภาพ
- คลาสเจ้าของรูปภาพ

ในการสืบค้นโดยใช้ “คำค้น” นั้น ระบบทำการสืบค้นคำค้นจากฐานข้อมูลรูปภาพ งานวิจัยนี้ได้กำหนดให้คำนวณค่าน้ำหนักของคำค้นที่ตรงกับคลาสต่างๆที่กำหนดไว้ โดยที่ผู้ใช้สามารถระบุเงื่อนไขการสืบค้นได้มากกว่า 1 เงื่อนไข แต่ระบบไม่สามารถตัดสินใจได้ว่าผู้ใช้งานระบบมีความสนใจ หรือมีความชอบเงื่อนไขใดมากกว่ากัน ดังนั้น งานวิจัยนี้จะไม่คำนวณค่าน้ำหนัก ในกรณีที่เป็นการสืบค้นโดยใช้เงื่อนไขการสืบค้น

ตัวอย่าง คำค้น : “Garfield cat”

1. ผู้ใช้งานระบบทำการสืบค้นข้อมูลรูปภาพ ด้วยการระบุ “คำค้น” ดังแสดงในภาพประกอบที่ 3.9



ภาพประกอบที่ 3-8 ภาพตัวอย่างหน้าจอรับคำค้น

2. ระบบนำคำค้นที่ผู้ใช้ป้อนเข้าในระบบ เปรียบเทียบกับคำศัพท์ในพจนานุกรมคำศัพท์เพื่อให้ระบบได้ลักษณะการสืบค้น จากคำค้นที่ผู้ใช้ป้อนเข้าระบบดังแสดงในข้อ 1 ตรงกันกับคำศัพท์ในพจนานุกรมคำศัพท์ประเภทที่ 5 คำอธิบายของรูปภาพ ดังนั้นระบบจะนำคำค้นนั้น มาผ่านกระบวนการตัดคำ โดยใช้โปรแกรมแบ่งคำ SWATH ทำให้ได้ผลลัพธ์จากการตัดคำ ดังนี้

“Cat” จากคุณสมบัติในภาษา OWL ของ Data Properties<Name>Cat</Name> (Data Properties: Name)

“Garfield” จากคุณสมบัติในภาษา OWL ของ Data Properties<Imagesearch:Featured>Garfield was originally created by Davis with the intention to come up with a good, marketable character</ Imagesearch:Featured > (Data Properties: Feature)

3. จากนั้น ระบบนำคำที่ตัดได้นี้ ไปเปรียบเทียบกับพจนานุกรมคำศัพท์ ประเภทเกี่ยวกับรูปภาพ และนำคำที่ตัดได้ที่ตรงกันกับคำศัพท์ในพจนานุกรมคำศัพท์ มาเข้าประโยคคิวรีด้วยภาษาสืบค้น SPARQL เพื่อทำการสืบค้นคำในฐานความรู้ออนโทโลยี

ขั้นตอนการพัฒนางานวิจัย

ระบบสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์ ประกอบด้วย 3 ขั้นตอน คือ การสร้างพจนานุกรมคำศัพท์ การออกแบบฐานความรู้ออนโทโลยีและการสืบค้นรูปภาพเชิงความหมาย

1. การสร้างพจนานุกรมคำศัพท์

เป็นการใช้ Text File บรรจุคำศัพท์จากเว็บไซต์ที่เกี่ยวข้องกับรูปภาพ เพื่อใช้เป็นฐานข้อมูลคำศัพท์ที่เกี่ยวข้องกับรูปภาพ สำหรับนำไปใช้ในการเปรียบเทียบกับคำค้นที่ผู้ใช้ป้อนเข้ามาในระบบ และเพื่อช่วยให้ระบบเข้าใจว่าจากคำค้นนั้นๆ ควรจะใช้ลักษณะการสืบค้นแบบใดเพื่อให้ผลลัพธ์ที่ได้จากการเปรียบเทียบมีประสิทธิภาพ และช่วยให้ใช้เวลาในการสืบค้นเร็วมากขึ้น

การออกแบบฐานความรู้ออนโทโลยี

เป็นการออกแบบออนโทโลยีรูปภาพที่ประกอบด้วยคลาสหลักต่างๆ และนำข้อมูลของรูปภาพมาบรรยายให้อยู่ในรูปแบบของ RDF โดยใช้ OWL ที่ได้มีการบรรยายลักษณะของรูปภาพตามโครงสร้างของออนโทโลยีที่ได้ออกแบบไว้ เพื่อรวบรวมไว้เป็นฐานความรู้สำหรับนำไปใช้ในการสืบค้นต่อไป ในการนิยามโครงสร้างการอธิบายข้อมูลของภาพ เป็นการอธิบายเพื่อให้มีลักษณะของภาษาเว็บเชิงความหมาย ซึ่งประกอบด้วยคุณสมบัติต่างๆ ทั้งคุณสมบัติของวัตถุ และคุณสมบัติของชนิดข้อมูล โดยที่ข้อมูลเหล่านี้ถูกแสดงด้วยเซตของทริปเปิล <subject, predicate, object>

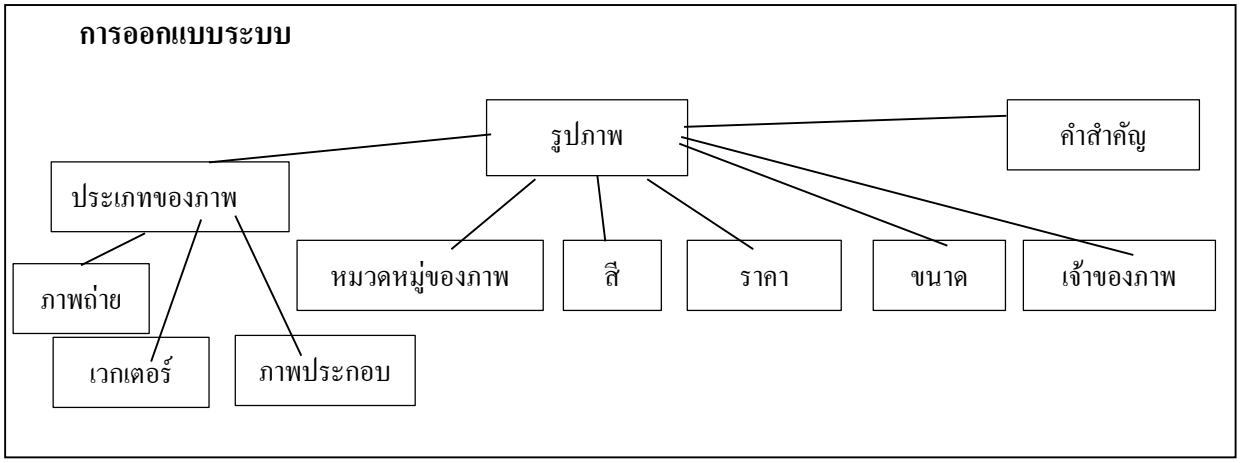
2. การสืบค้นรูปภาพเชิงความหมาย

ขั้นตอนนี้ ระบบกำหนดให้ผู้ใช้งานระบบสามารถสืบค้นข้อมูลที่ต้องการ ด้วยการระบุคำค้นที่ต้องการ หรือระบุเงื่อนไขที่ต้องการสืบค้นข้อมูลที่ต้องการ เพื่อให้ระบบนำไปประมวลผลตามขั้นตอนการทำงาน 4 ขั้นตอน คือ การเลือกคำสำคัญ การเลือกกลุ่มคำสืบค้น การสืบค้นในฐานความรู้ออนโทโลยี และการคำนวณค่าน้ำหนักของผลลัพธ์ที่ได้จากการสืบค้นข้อมูลรูปภาพ เพื่อให้ระบบแสดงผลลัพธ์จากการสืบค้นได้ตรงตามความต้องการของผู้ใช้งานระบบ

ระยะเวลาในการดำเนินการศึกษา

ตารางที่ 3-6 แผนการดำเนินงาน

ลำดับ	ขั้นตอนการดำเนินงาน	ระยะเวลาในการดำเนินการ															
		สิงหาคม				กันยายน				ตุลาคม				พฤศจิกายน			
		ปี 2556				ปี 2556				ปี 2556				ปี 2556			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	การศึกษาเบื้องต้น																
2	การวิเคราะห์ระบบ																
3	การออกแบบระบบ																
4	การติดตั้งทดสอบโปรแกรม																
5	การแก้ไขโปรแกรม																
6	เขียนเอกสารโครงการ																



ภาพประกอบที่ 3-9 กรอบแนวคิดในการวิจัยการพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

ตามกรอบแนวคิดภาพประกอบที่ 3-9 การพัฒนาออนโทโลยีสำหรับการสืบค้นรูปภาพเชิงความหมาย โดยการจำแนกรูปภาพในลักษณะข้อมูลรายละเอียดที่อธิบายถึงความเป็นมาของข้อมูล (Metadata) (Jia Liu,2007) เพื่อให้ได้ผลลัพธ์จากการสืบค้นรูปภาพตรงตามความต้องการของผู้ใช้งาน

ประเภทของภาพ คือได้จำแนกจากนามสกุลของภาพและลักษณะทางเทคนิคจำแนกได้ คือ ภาพถ่าย ภาพเวกเตอร์ และภาพประกอบ

หมวดหมู่ของภาพ คือได้จำแนกจากลักษณะองค์ประกอบและความหมายในภาพนั้น ๆ เช่น ภาพความรัก ภาพธรรมชาติ ภาพสัตว์ เป็นต้น

สี คือโทนสีของภาพ

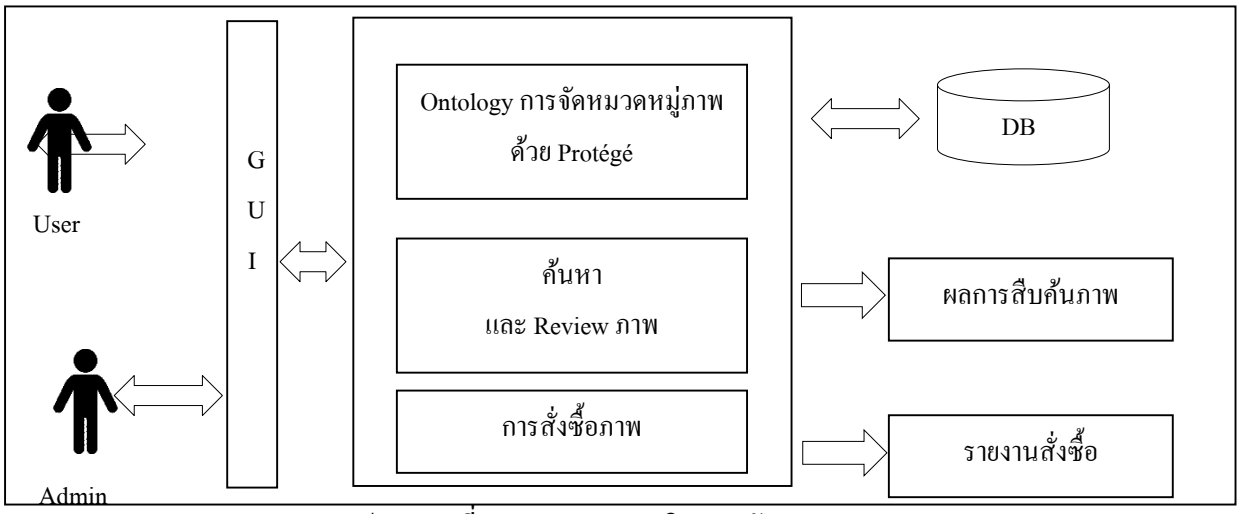
ราคา คือราคาในการขายภาพแต่ละภาพ

ขนาด คือขนาดของภาพ

เจ้าของภาพ คือผู้ที่ขายภาพในระบบ

คำสำคัญ คือคำอธิบายภาพที่เจ้าของภาพกำหนด

คำสำคัญ คือคำอธิบายภาพที่เจ้าของภาพกำหนด



ภาพประกอบที่ 3-10 กรอบแนวคิดการพัฒนา
ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

ตามกรอบแนวคิดภาพประกอบที่ 3-10 ระบบประกอบด้วยหน่วยที่สร้างออนโทโลยีสำหรับรูปภาพและมีการเก็บรูปภาพในฐานข้อมูลเพื่อให้บริการ User

User คือผู้ที่ต้องการสืบค้นและสั่งซื้อรูปภาพที่ต้องการผ่าน GUI

Admin คือผู้ที่ทำการปรับปรุงข้อมูลรูปภาพที่ต้องการจะขายเก็บไว้ในฐานข้อมูลและปรับปรุงข้อมูลหมวดหมู่และลักษณะของออนโทโลยี

วิเคราะห์การนำออนโทโลยีมาใช้

ได้ทำการศึกษาข้อมูลเพิ่มเติมเกี่ยวกับแนวทางการนำออนโทโลยีมาช่วยในการจำแนกประเภทของรูปภาพเพื่อค้นหาให้ตรงตามความต้องการ โดยทำการวิเคราะห์ข้อมูลเกี่ยวกับประเภทและคุณลักษณะของรูปภาพนั้นๆ ประเภท อารมณ์ ความหมาย และลักษณะทางกายภาพ จากนั้นจึงออกแบบและจัดเก็บข้อมูลลงบนออนโทโลยีเพื่อให้ได้ข้อมูลที่มีความสัมพันธ์กันอย่างถูกต้อง ซึ่งในการค้นหาจะช่วยให้ได้ผลลัพท์ที่ตรงใจผู้ใช้มากที่สุด การจัดเก็บข้อมูลลงบนออนโทโลยี มีความสะดวก ข้อมูลมีความสัมพันธ์กันจริง ผลลัพท์ในการทดสอบการค้นหาผ่านโปรแกรม Protégé มีความถูกต้องแต่ยังไม่สามารถนำมา ประยุกต์ใช้กับเว็บไซต์ได้อย่างเต็มที่ทำได้เพียงการเก็บเป็นฐานข้อมูลและหากมีการพัฒนาต่อ โดยเพิ่มข้อมูลให้ครอบคลุมคุณลักษณะอื่น ๆเพิ่มเติม น่าจะทำให้ออนโทโลยีสามารถประมวลผลได้อย่างหลากหลายและมีประสิทธิภาพมากขึ้นกว่าเดิม

บทที่ 4

ผลการศึกษา

งานวิจัยนี้ ได้พัฒนาวิธีการสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยี โดยผู้วิจัยได้ใช้ Protégé 5.0.1 ในการสร้างฐานความรู้ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์ และใช้ภาษา SPARQL ในการเขียนคิวรี (Query) ข้อมูลจากฐานความรู้ออนโทโลยี ทำให้ได้วิธีการสืบค้นรูปภาพที่มีประสิทธิภาพสามารถสืบค้นข้อมูลได้ตรงตามความต้องการของผู้ใช้งาน

การทดลองและข้อกำหนดเบื้องต้นของการทดลอง

ผู้วิจัยได้ทำ การทดลองเพื่อทดสอบและประเมินประสิทธิภาพการทำ งานของระบบด้วยการประเมินความถูกต้องของการสืบค้นข้อมูลรูปภาพและประเมินความพึงพอใจของผู้ใช้งานระบบในการทดลอง ผู้วิจัยได้กำหนดข้อกำหนดเบื้องต้น ดังนี้

1. ข้อมูลรูปภาพและคำค้น

ผู้วิจัยได้ทำการสุ่มรูปภาพจากฐานข้อมูลจำนวน 200 รูป รูปภาพเหล่านี้ ถูกบรรยายและจัดเก็บข้อมูลรูปภาพอยู่ในรูปแบบของ RDF (Resource Description Framework) โดยใช้ OWL (Ontology Web Langrage) ที่ได้มีการจำแนกรูปภาพตามโครงสร้างของออนโทโลยีที่ได้ออกแบบไว้ และให้ผู้ใช้งานระบบจำนวน 35 คน ทำการสืบค้นข้อมูลรูปภาพ ด้วยการระบุคำค้นที่แตกต่างกันจำนวน 20 คำ ดังแสดงในตารางที่ 4-1

ตารางที่ 4-1 คำและวลีสืบค้นที่ใช้การทดลองออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

1. Animal	5. Dog	9. Yellow	13. Green	17. Tree
2. Flower	6. Smile	10. Sun	14. Football	18. Car
3. Park	7. Love	11. Baby	15. Moon	19. Food
4. Wedding	8. Fashion	12. Vectors	16. Underwater	20. Somtom

2. การวัดความถูกต้องของการสืบค้นข้อมูลรูปภาพ

ผู้วิจัยได้ทำการวัดความถูกต้องของการสืบค้นข้อมูลรูปภาพ โดยใช้สมการวัดความแม่นยำ (Precision) และการเรียกซ้ำ (Recall) ของการสืบค้นแต่ละคำหลัก เพื่อหาค่าเฉลี่ยของการสืบค้น โดยมีเกณฑ์การประเมินดังตารางที่ 4.2

$$\text{Recall} = \frac{TruePositive}{TruePositive+FalseNegative}$$

เมื่อ TruePositive หมายถึง จำนวนข้อมูลที่ทำนายถูกว่าเป็นคลาสซึ่งกำลังสนใจอยู่

FaleseNegative หมายถึง จำนวนข้อมูลที่ทำนายผิดมาเป็นคลาสซึ่งไม่ได้สนใจอยู่

หาค่าการแม่นยำของโมเดล (Precision) (เอกสิทธิ์ พัชรวงศ์ศักดา, 2557) โดยพิจารณาแยกทีละคลาส จากสูตร

$$\text{Precision} = \frac{TruePositive}{TruePositive+FalsePositive}$$

เมื่อ TrueNegative หมายถึง จำนวนข้อมูลที่ทำนายถูกว่าเป็นคลาสซึ่งไม่ได้สนใจอยู่

FalsePositive หมายถึง จำนวนข้อมูลที่ทำนายผิดมาเป็นคลาสซึ่งกำลังสนใจอยู่

3. การวิเคราะห์ข้อมูล

งานวิจัยนี้กำหนดให้ใช้แบบสอบถาม ในการสอบถามความพึงพอใจของผู้ใช้งานระบบและทำการวิเคราะห์ข้อมูลที่ได้จากแบบสอบถาม เพื่อประเมินความพึงพอใจของผู้ใช้งานระบบโดยหาค่าความสอดคล้อง (Item-Objective Congruence: IOC) (ไพศาล วรคำ, 2555) จากสูตร

$$\text{IOC} = \frac{\sum R}{N}$$

เมื่อ IOC คือ ความสอดคล้องระหว่างวัตถุประสงค์กับแบบทดสอบ

$\sum$R หมายถึง ผลรวมของความเห็นของผู้เชี่ยวชาญ

N หมายถึง จำนวนผู้เชี่ยวชาญที่ประเมินความสอดคล้องในข้อนั้น

ตารางที่ 4-2 เกณฑ์การแปลความหมายข้อมูลและพิจารณา จากค่า Precision และ Recall ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์

ระดับเกณฑ์	ความหมาย
0.85-1.00	ประสิทธิภาพของระบบอยู่ในระดับดีมาก
0.75-0.84	ประสิทธิภาพของระบบอยู่ในระดับดี
0.55-0.74	ประสิทธิภาพของระบบอยู่ในระดับปานกลาง
0.35-0.54	ประสิทธิภาพของระบบอยู่ในระดับน้อย
0.00-0.34	ประสิทธิภาพของระบบอยู่ในระดับใช้ไม่ได้

การทดสอบความถูกต้องของการสืบค้นข้อมูลรูปภาพ

งานวิจัยนี้ได้ทำการทดสอบความถูกต้องของการสืบค้นข้อมูลรูปภาพ ด้วยการทดลองสืบค้นข้อมูลรูปภาพที่ให้ผู้ใช้งานระบบระบุคำค้น หรือเงื่อนไขการสืบค้น เพื่อให้ระบบทำการสืบค้นข้อมูลรูปภาพตามที่ผู้ใช้งานระบบต้องการ

1. การทดสอบความถูกต้องของการสืบค้นข้อมูลรูปภาพด้วยการระบุคำค้น

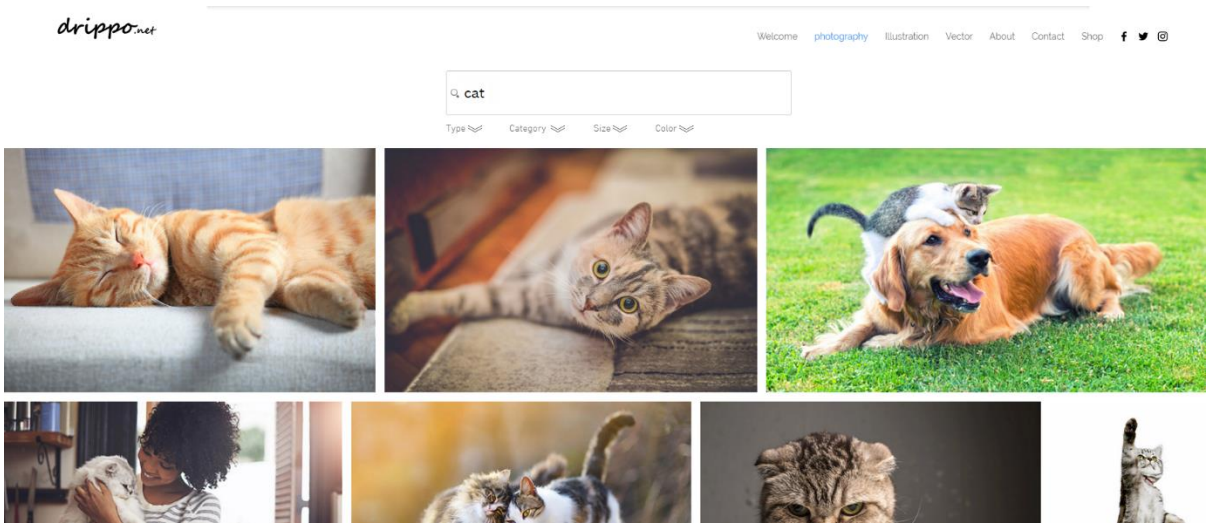
ผู้วิจัยได้ทำการทดลองสืบค้นข้อมูลรูปภาพด้วยการระบุคำค้น เพื่อทดสอบความถูกต้องของการสืบค้นข้อมูลรูปภาพ โดยได้แบ่งคำค้นข้อมูลรูปภาพเป็น 1 กรณี ดังนี้

- กรณีที่ 1: คำค้นเป็นรายละเอียดของรูปภาพ

ผลการประเมินความถูกต้องของการสืบค้นข้อมูลรูปภาพด้วยการระบุคำค้นจะเห็นได้ว่า มีค่าเฉลี่ยของ Precision เท่ากับ 87% และค่าเฉลี่ยของ Recall เท่ากับ 100%

2. การประเมินความถูกต้องของการสืบค้นข้อมูลรูปภาพด้วยการระบุเงื่อนไขการสืบค้น

การสืบค้นข้อมูลรูปภาพด้วยการระบุเงื่อนไขการสืบค้นนั้น ผู้ใช้งานระบบอาจจะระบุคำค้นที่เป็นคำอธิบายของรูปภาพ หรือระบุลักษณะและประเภทของรูปภาพ โดยที่ผู้ใช้งานระบบสามารถระบุลักษณะและประเภทของรูปภาพได้



ภาพประกอบที่ 4-1 ตัวอย่างการสืบค้นด้วยการระบุเงื่อนไขการสืบค้น

ผู้วิจัยได้ทำการทดลองสืบค้นข้อมูลรูปภาพ เพื่อทดสอบความถูกต้องของการสืบค้นข้อมูลรูปภาพด้วยการระบุเงื่อนไขการสืบค้น โดยได้แบ่งเงื่อนไขการสืบค้นข้อมูลรูปภาพเป็น 3 รูปแบบ ดังนี้

3. สรุปผลการทดสอบความถูกต้องของการสืบค้นข้อมูลรูปภาพ

จากการทดลองพบว่าการทดสอบความถูกต้องของการสืบค้นรูปภาพ จากการพัฒนาระบบการสืบค้นรูปภาพเชิงความหมาย มีระดับ Precision เฉลี่ยเท่ากับ 87% ซึ่งหมายถึงระบบมีประสิทธิภาพในระดับดี ในขณะที่ระดับ Recall มีประสิทธิภาพเฉลี่ยอยู่ที่ 100% ซึ่งมีประสิทธิภาพในระดับดีมาก

การประเมินผล

การประเมินความพึงพอใจของผู้ใช้งานระบบ

ผู้วิจัยได้ทำการประเมินความพึงพอใจของผู้ใช้งานระบบ ด้วยการกำหนดให้ผู้ใช้งานระบบจำนวน 25 คน ทำแบบสอบถามเพื่อประเมินประสิทธิภาพการทำงานของระบบใน 3 ด้าน คือ

- ด้านความสัมพันธ์ และการแบ่งหมวดหมู่ของข้อมูล
- ด้านการใช้งานของโปรแกรม
- ด้านการแสดงผลของโปรแกรม

ผู้วิจัยได้กำหนดให้ใช้เกณฑ์การให้คะนนในแบบประเมิน และใช้สถิติเชิงพรรณนาในการวัดค่ากลางโดยใช้ค่าตัวกลางเลขคณิต และวัดการกระจายของข้อมูลโดยใช้ส่วนเบี่ยงเบนมาตรฐาน โดยมีเกณฑ์การแปลความหมายข้อมูลและพิจารณาจากค่าตัวกลางเลขคณิต ดังแสดงในตารางที่ 4-3

1. ผลการประเมินความพึงพอใจด้านความสัมพันธ์ และการแบ่งหมวดหมู่ของข้อมูล

ผู้วิจัยใช้วิธีค้นคว้าจากแหล่งข้อมูลต่างๆและตกผลึกเป็นข้อมูล จากนั้นได้ทำการออกแบบสอบถามให้ตรงตามข้อมูล เพื่อทดสอบความเที่ยงตรงตามเนื้อหา โดยดำเนินการให้ผู้เชี่ยวชาญในด้านเนื้อหาจำนวน 5 ท่าน พิจารณาถึงความสอดคล้องระหว่างวัตถุประสงค์กับแบบทดสอบโดยพิจารณาเป็นรายข้อ วิธีการพิจารณานี้เรียกว่า การหาค่าสัมประสิทธิ์ความสอดคล้อง (Index of Item – Objective Congruence : IOC) (ไพศาล วรคำ, 2555) โดยได้ค่าเฉลี่ยจากคำตอบของผู้เชี่ยวชาญด้านเนื้อหา ดังตารางที่ 4-3

ตารางที่ 4-3 แสดงค่าความตรงตามเนื้อหา (IOC) ของข้อมูล

ข้อที่	รายละเอียด	ระดับความสอดคล้อง		
		สอดคล้อง (1)	ไม่แน่ใจ (0)	ไม่สอดคล้อง (-1)
1	ลักษณะของภาพมีความสำคัญต่อการจำแนกออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์	1	-	-
2	ขนาดของภาพมีความสำคัญต่อการจำแนกออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์	1	-	-
3	ประเภทของภาพมีความสำคัญต่อการจำแนกออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์	1	-	-
4	ราคาของภาพมีความสำคัญต่อการจำแนกออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์	1	-	-
5	สีของภาพมีความสำคัญต่อการจำแนกออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์	1	-	-
6	คำสำคัญของภาพมีความสำคัญต่อการจำแนกออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์	1	-	-
7	เจ้าของของภาพมีความสำคัญต่อการจำแนกออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์	1	-	-
ค่า IOC เฉลี่ย		**1**		

จากตารางที่ 4-3 การตรวจสอบค่าความเที่ยงตรงด้านเนื้อหาสามารถกระทำโดยนำแบบทดสอบให้ผู้เชี่ยวชาญพิจารณาว่า แบบสอบถามแต่ละข้อมีความสอดคล้องกับวัตถุประสงค์เชิงพฤติกรรมหรือไม่อย่างไร ถ้ามีความสอดคล้องผู้เชี่ยวชาญจะให้ค่าเป็น “+1” แต่ถ้าผู้เชี่ยวชาญเห็นว่าข้อสอบข้อนั้นไม่มีความสอดคล้องกับวัตถุประสงค์จะให้ค่าเป็น “-1” และในกรณีที่ผู้เชี่ยวชาญไม่แน่ใจว่าข้อสอบข้อนั้นมีความสอดคล้องกับวัตถุประสงค์หรือไม่ก็จะให้ค่าเป็น “0” (ไพศาล วรคำ, 2555)

สรุปการประเมินประสิทธิภาพการทำงานของระบบ

ผู้วิจัยทำการทดลอง 20 ข้อมูลสืบค้นที่แตกต่างกัน มี 3 คำสืบค้นที่ไม่มีผลลัพธ์ ส่วนอีก 17 คำสืบค้นมีผลลัพธ์ ผลการวัดประสิทธิภาพในการสืบค้นด้วยวิธี Recall เท่ากับ 0.85 และวิธี Precision เท่ากับ 0.82 ผลการทดลองที่ได้คือ ประสิทธิภาพความถูกต้องของระบบอยู่ที่ระดับระดับดีมาก และประสิทธิภาพความแม่นยำของระบบอยู่ที่ระดับระดับดี ตามเกณฑ์การแปลความหมายของข้อมูลที่ได้จากการประเมิน ระบบเข้าใจความหมายของคำที่มีความหมายเหมือนกัน ทำให้ผลลัพธ์ตรงกับความต้องการของผู้ใช้และผลลัพธ์ที่ได้ก็มีความครบถ้วนมากขึ้น เพราะการบรรยายภาพของแต่ละผู้ใช้อาจจะใช้คำที่แตกต่างกัน แต่ในเชิงความหมายมีความหมายเหมือนกัน

จากการทดลองพบว่าการทดสอบความถูกต้องของการสืบค้นรูปภาพออนโทโลยีชุดนี้สอดคล้องกับวิธีการสืบค้นรูปภาพเชิงความหมายทำให้ได้วิธีการสืบค้นข้อมูลรูปภาพที่มีประสิทธิภาพสามารถสืบค้นข้อมูลได้ถูกต้องเป็นที่ยอมรับได้ ซึ่งสอดคล้องกับงานวิจัยเรื่อง การพัฒนาออนโทโลยีสำหรับระบบให้คำแนะนำการบริโภคอาหาร ตามโภชนาการเฉพาะบุคคล (นภัส สุขสมและคณะ, 2559) ด้านการพัฒนาออนโทโลยี และงานวิจัยเรื่องการสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยี (อรวรรณ อุไรเรืองพันธและสมจิตร อาจอินทร์, 2552) ด้านการประเมินประสิทธิภาพของออนโทโลยีจากการวิจัยครั้งนี้ผลการวัดประสิทธิภาพในการสืบค้นด้วยวิธี Recall เท่ากับ 0.85 และวิธี Precision เท่ากับ 0.82 ผลการทดลองที่ได้คือ ประสิทธิภาพความถูกต้องของระบบอยู่ที่ระดับระดับดีมาก และประสิทธิภาพความแม่นยำของระบบอยู่ที่ระดับระดับดี

บทที่ 5

สรุปผลการทดลองและข้อเสนอแนะ

สรุปและอภิปรายผล

สรุปผลการดำเนินงาน

ระบบการสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยี ได้พัฒนาวิธีการสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์ โดยเป็นการพัฒนาออนโทโลยีที่สอดคล้องกับวิธีการสืบค้นรูปภาพเชิงความหมาย ทำให้ได้วิธีการสืบค้นรูปภาพที่มีประสิทธิภาพ สามารถสืบค้นข้อมูลได้ถูกต้อง ครบถ้วนตรงตามความต้องการของผู้ใช้งานระบบ โดยที่ระบบมีความสามารถในการทำงานดังต่อไปนี้

1. ระบบสามารถจัดการฐานความรู้ออนโทโลยีของระบบ โดยสามารถเพิ่ม ลบ แก้ไขข้อมูลรูปภาพ

2. ระบบสามารถรับคำค้นจากการคิวรีของผู้ใช้ที่เป็นคำค้น เพื่อนำไปใช้ในการสืบค้นข้อมูลรูปภาพตามความต้องการของผู้ใช้

3. จากการทดสอบการใช้งานระบบโดยผู้ใช้งานระบบจำนวน 35 คน ทำการสืบค้นข้อมูลรูปภาพจากฐานข้อมูลจำนวน 200 รูป ด้วยการระบุคำค้นที่แตกต่างกันจำนวน 20 คำ

- ระบบมีความถูกต้องของการสืบค้นข้อมูลรูปภาพ จากการทดลองพบว่าการทดสอบความถูกต้องของการสืบค้นรูปภาพ จากการพัฒนาระบบการสืบค้นรูปภาพเชิงความหมาย มีระดับ Precision เฉลี่ยเท่ากับ 82% ซึ่งหมายถึงระบบมีประสิทธิภาพในระดับดี ในขณะที่ระดับ Recall มีประสิทธิภาพเฉลี่ยอยู่ที่ 85% ซึ่งมีประสิทธิภาพในระดับดีมาก

- ผู้ใช้งานระบบมีความพึงพอใจในการใช้งานระบบด้านต่างๆ ดังนี้

  - ด้านความสัมพันธ์และการแบ่งหมวดหมู่ของข้อมูลผู้ใช้งานระบบมีความพึงพอใจระดับดี

  - ด้านผลการประเมินด้านการใช้งานโปรแกรมอยู่ ผู้ใช้งานระบบมีความพึงพอใจระดับดี

- ด้านผลการประเมินด้านการแสดงผลของโปรแกรมอยู่ ผู้ใช้งานระบบมีความพึงพอใจระดับดี

ปัญหาอุปสรรค

ผู้วิจัยได้พัฒนาระบบการจัดเก็บข้อมูลรูปแบบสำหรับการสืบค้นข้อมูลจากการนำภาพขึ้นระบบของผู้ขาย นอกจากนั้นการเขียนคำบรรยายลักษณะภาพของผู้เขียนแต่ละคนอาจมีการใช้คำบรรยายที่แตกต่างกันแต่มีความหมายเหมือนกัน หากระบบการสืบค้นข้อมูลใช้หลักการสืบค้นข้อมูลด้วยคำสำคัญระบบจะแสดงผลการสืบค้นข้อมูล ด้วยการแสดงเฉพาะเอกสารที่ตรงกันกับคำที่ใช้สืบค้นเท่านั้น

การออกแบบระบบการสืบค้นข้อมูลให้มีการใช้คำที่มีความหมายเหมือนกัน (Synonym) จะทำให้คอมพิวเตอร์สามารถเข้าใจถึงกลุ่มคำที่มีความหมายเหมือนกัน และนำมาใช้ในการสืบค้นข้อมูลผลลัพธ์ที่ได้จึงเป็นผลลัพธ์ ในเชิงความหมายที่มีความแม่นยำสูงกว่าการสืบค้นข้อมูลด้วยคำสำคัญ (Keyword-Based) ช่วยเพิ่มประสิทธิภาพในการสืบค้นข้อมูลรูปภาพ ทำให้ผู้ใช้งานระบบสามารถสืบค้นข้อมูลได้ถูกต้อง ครบถ้วน ตรงตามความต้องการมากกว่าการสืบค้นข้อมูลในปัจจุบันที่เป็นการสืบค้นด้วยคำสำคัญ (Keywords)

ข้อเสนอแนะ

งานวิจัยนี้ได้สร้างตัวแบบออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์ เพื่อให้ระบบนำไปใช้ในการสืบค้นข้อมูลรูปภาพ ในการสร้างตัวแบบนั้น ผู้วิจัยได้ใช้ข้อมูลรูปภาพที่มีคำอธิบายรูปภาพ ทำให้เป็นข้อจำกัดในการใช้งานระบบที่ผู้ใช้งานสามารถสืบค้นได้เฉพาะข้อมูลรูปภาพที่ผู้ชายขึ้นระบบไว้เท่านั้น ในการพัฒนาต่อไปนั้น ผู้พัฒนาอาจพิจารณาพัฒนาออนโทโลยีที่บรรยายข้อมูลรูปภาพต่างๆจากทั่วโลก รวมถึงการเชื่อมโยงข้อมูลที่เกี่ยวข้องกับรูปภาพที่ทำการสืบค้นออกมาได้ และพัฒนาออนโทโลยีให้เหมาะสำหรับใช้กับระบบการสืบค้นรูปภาพสำหรับธุรกิจขายภาพออนไลน์เนื่องจากเหมาะกับยุคการพัฒนาประเทศแบบ ไทยแลนด์ 4.0 หรือยุคการนำระบบดิจิทัลเพื่อพัฒนาเศรษฐกิจและสังคม

ฐานข้อมูลรูปภาพที่ใช้ในงานวิจัยนี้ รองรับเฉพาะคำสำคัญที่ผู้วิจัยจัดเก็บและแยกประเภทในลักษณะของหมวดหมู่ ลำดับชั้น คำที่เขียนเหมือนกันแต่ความหมายต่างกันคำที่เขียนต่างกันแต่มีความหมายเหมือนกัน และคำที่มีความสัมพันธ์ระหว่างกันได้เอง โดยที่การทำงานในขั้นตอนการนำเข้าข้อมูล การคัดแยกข้อมูล และการออกแบบโครงสร้างในฐานข้อมูลออนโทโลยีรูปภาพเป็นการประมวลผลด้วยมือ (Manual system) ทำให้มีโอกาสเกิดข้อผิดพลาด เสียเวลาในการประมวลผลข้อมูลที่รวบรวมไว้ในฐานข้อมูลอาจจะไม่ครอบคลุมการใช้งาน ทำให้ผู้ใช้งานระบบไม่ได้รับข้อมูลที่ตรงตามความต้องการเท่าที่ควรในกรณีที่ต้องการสืบค้นในลักษณะที่ลงรายละเอียดมากขึ้น ในการพัฒนางานวิจัยต่อไปนั้น ผู้พัฒนาควรพิจารณานำ WordNet และ Thesaurus เข้ามาใช้ในการนิยามออนโทโลยี และทำการหาค่าความคล้ายคลึงของคำสืบค้นโดยอาศัยความสัมพันธ์ของความหมายใน Wordnet เพื่อให้ได้ข้อมูลรูปภาพประกอบที่มีความสัมพันธ์ในเชิงความหมายและตรงตามความต้องการของผู้ใช้งานระบบ โดยการสืบค้นรูปภาพสำหรับธุรกิจขายภาพออนไลน์นี้จัดทำเพื่อเป็นตัวแทนระบบร้านค้าขายรูปภาพซึ่งเมื่อกาลเวลาเปลี่ยนแปลงจะทำให้ระบบดังกล่าวด้อยลงและจะมีประสิทธิภาพมากขึ้นจะต้องมีการปรับปรุงหรือพัฒนางานวิจัยต่อยอดต่อไป

นอกจากนี้ผู้พัฒนาควรพัฒนาให้ระบบมีความสามารถปรับปรุงฐานข้อมูลแบบอัตโนมัติเพื่อให้ฐานข้อมูลของระบบมีความทันสมัยอยู่เสมอ และพัฒนาเพื่อให้ได้ระบบการสืบค้น ที่ได้แก้ปัญหาในเรื่องที่ผู้ใช้ป้อนคำสืบค้นผิด หรือป้อนคำสืบค้นที่ไม่ถูกต้องตามหลักภาษาไทยและภาษาอังกฤษและเป็นระบบการสืบค้นที่มีเครื่องมือช่วยแนะนำคำสืบค้น เพื่อเป็นการเพิ่มความสะดวกให้ผู้ใช้งานระบบในการป้อนคำสืบค้นหรือวลีสืบค้น

งานวิจัยนี้ผู้วิจัยใช้วิธีการ Weight Hierarchies ที่กำหนดให้ระบบคำนวณค่าน้ำหนักของผลลัพธ์ที่ได้จากการสืบค้นข้อมูลรูปภาพ เพื่อให้ระบบนำค่าน้ำหนักที่ได้ไปใช้ในการจัดลำดับการแสดงผลลัพธ์ที่ได้จากการสืบค้น โดยที่งานวิจัยนี้ไม่ได้มีการพิจารณากรณีที่เป็นการสืบค้นข้อมูลรูปภาพที่มีรายละเอียดของรูปภาพมาก และมีข้อมูลบางส่วนของรูปภาพตรงกันกับคำสืบค้นทำให้มีความคลาดเคลื่อนในการแสดงผลลัพธ์ ในการพัฒนางานวิจัยต่อไป ผู้พัฒนาควรพิจารณาให้ระบบคำนวณค่าน้ำหนักในกรณีที่รายละเอียดของรูปภาพไม่ตรงกันกับคำสืบค้น เพื่อให้ได้ผลลัพธ์ของการสืบค้นที่ถูกต้อง และตรงตามความต้องการของผู้ใช้มากที่สุด

บรรณานุกรม

ธีรวิชญ์ วงษา. (2557). ออนโทโลยีกับการจัดการความรู้. , มหาวิทยาลัยเชียงใหม่: ข่าวสารคณะวิทยาศาสตร์ มช.

นฤพนธ์ พนาวงศ์. (2553). การพัฒนาระบบสืบค้นข้อมูลออนโทโลจีท่องเที่ยวด้วยภาษา SPARQL., มหาวิทยาลัยนเรศวร: Proceedings การประชุมวิชาการ “นเรศวรวิจัย” ครั้งที่ 7.

ไพศาล วรคำ. (2555). การวิจัยทางการศึกษา. , มหาสารคาม: ตักสิลาการพิมพ์.

มาลี กาบมาลา. (2549). ออนโทโลยี: แนวคิดการพัฒนา. , มหาวิทยาลัยขอนแก่น: บรรณารักษศาสตร์และสารนิเทศศาสตร์ มข. ปีที่ 24

สมชาย ปราการเจริญ. (2548). ออนโทโลยีทางเลือกของการพัฒนาฐานความรู้ในรูปแบบเชิงเนื้อหา. เอกสารประกอบการสัมนาทางวิชาการ The 5 th National Conference on Computing and Information Technology-NCCIT 2009. (หน้า 92-99). กรุงเทพฯ: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

เอกสิทธิ์ พัชรวงศ์ศักดา. (2557). การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้าไมน์นิงเบื้องต้น. , กรุงเทพมหานคร: บริษัท เอเชียดิจิตอลการพิมพ์ จำกัด.

อรวรรณ อุไรเรืองพันธ สมจิตรอาจอินทร์.(2552). การสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยี The 5th National Conference on Computing and Information Technology.

Hyvönen, E., Styrman, A., & Saarela, S. (2002). Ontology-based image retrieval. Proceedings of XML Finland 2002 Conference, 15-27.

Jia Liu. (2007). Metadata and Its Applications in the Digital Library: Approaches and Practices. , Westport, Connecticut: Libraries Unlimited.

Merriam-Webster’s Online Dictionary. (2004). Springfield, Mass.: Merriam-Webster. Mimoune, Mourad El-Hadj: Pierra, Gay and Ait-Ameur, Yemine.

Protégé Project (Copyright 2016). “What is protégé?.”. , สืบค้นวันที่ 25 กรกฎาคม 2559 จาก http://protege.stanford.edu/index.html/.

Stanford. (2014). “What is an Ontology?”. , สืบค้นวันที่ 14 สิงหาคม 2559 จาก http://wwwksl.stanford.edu/kst/what-is-an-ontology.html.

Shutterstock (2014). คู่มือความสำเร็จสำหรับช่างภาพ. , มหาวิทยาลัยราชภัฏธนบุรี: คณะวิทยาการจัดการ

บรรณานุกรม (ต่อ)

Vijay Walunj. (2008). Cloud-Based Semantic Image Search Engine. B.E. University of Mumbai, India.

Webneel. (2016). “Different Types of Photography.”. , สืบค้นวันที่ 14 สิงหาคม 2559 จาก http://webneel.com/different-types-of-photography.

ภาคผนวก

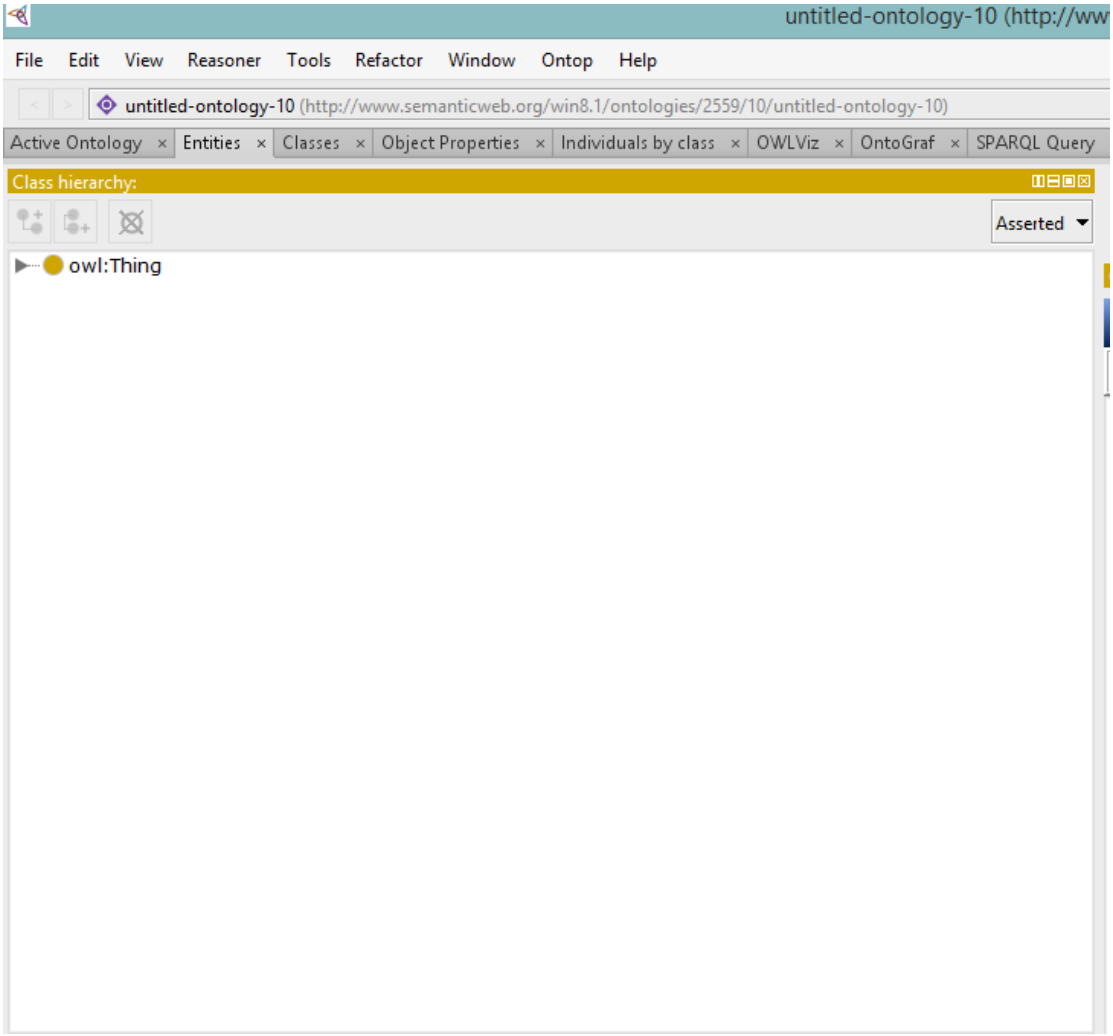
ภาคผนวก ก

คู่มือการใช้งาน

การสร้างออนโทโลยีโดยใช้โปรแกรม Protégé

1. โปรแกรม Protégé

Protégé เป็นซอฟต์แวร์ประเภทโอเพ่นซอร์ส ใช้งานสำหรับการสร้างออนโทโลยีและฐานความรู้ รองรับการสร้างออนโทโลยีด้วยภาษา RDF/RDFS และ OWL พัฒนา ขึ้นโดยมหาวิทยาลัยสแตนฟอร์ด ประเทศสหรัฐอเมริกา มีส่วนการติดต่อ ผู้ใช้งานเป็นแบบกราฟิก (Graphical User Interface: GUI)



ภาพประกอบที่ ผ-1 หน้าจอโปรแกรม Protége

2. การสร้าง OWL ชื่อ Tourism.owl

ออนโทโลยีรูปภาพที่ได้มีการออกแบบไว้ในบทที่ 3 จะนำการสร้าง Ontology ที่เป็นเอกสาร OWL โดยมีรายละเอียดของคลาสดังต่อไปนี้

โดย Class ประกอบด้วย

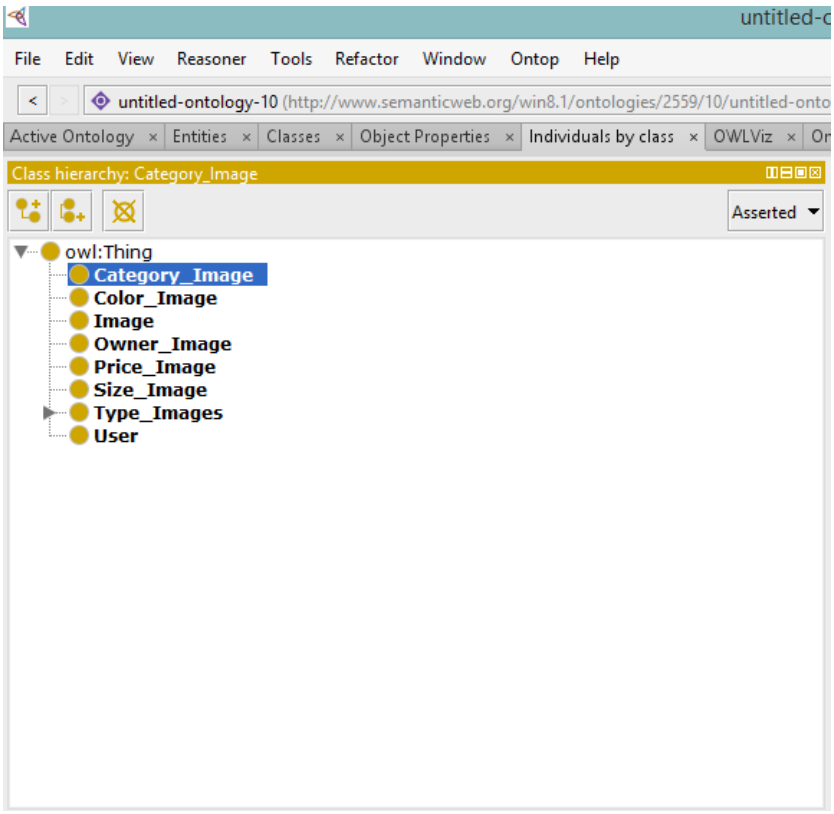
- Categoty
    - Abstract
    - Advertisement
    - Aerial
    - Amazing
    - Animals
    - Arts
    - Backgrounds
    - BeautifulWomen
    - Beauty
    - Bird
    - BlackAndWhite
    - Buildings
    - Business
    - Candid
    - Cat
    - ClipArt
    - Colorful
    - Dog
    - Drink
    - Editorial
    - Emotion
    - Fashion
    - Finance
    - Firework
    - Flower

- Food
- ForcedPerspective
- GlamourFashion
- HDR
- HiSpeed
- Icons
- Infographics
- Infrared
- Interiors
- Landmarks
- Landscape
- Love
- Macro
- ModelReleasedOnly
- Motion
- Nature
- NewBorn
- Night
- Outdoor
- Panorama
- Park
- PastAndPresent
- Patterns
- People
- Rainbow
- RealEstate
- Reflection
- Silhouette
- Sport
- StillLife

    - Storm
    - Sunrise
    - Textures
    - Tillshift
    - Timelapse
    - Travel
    - UnbanAndCity
    - Underwater
    - Vintage
    - Wedding
    - Wildlife
- Color
    - Any
    - Black
    - Blue
    - Brown
    - BW
    - DarkBlue
    - Full
    - Gray
    - Green
    - Orange
    - Pink
    - Purple
    - Red
    - Transparent
    - White
    - Yellow
- Image
- Owner

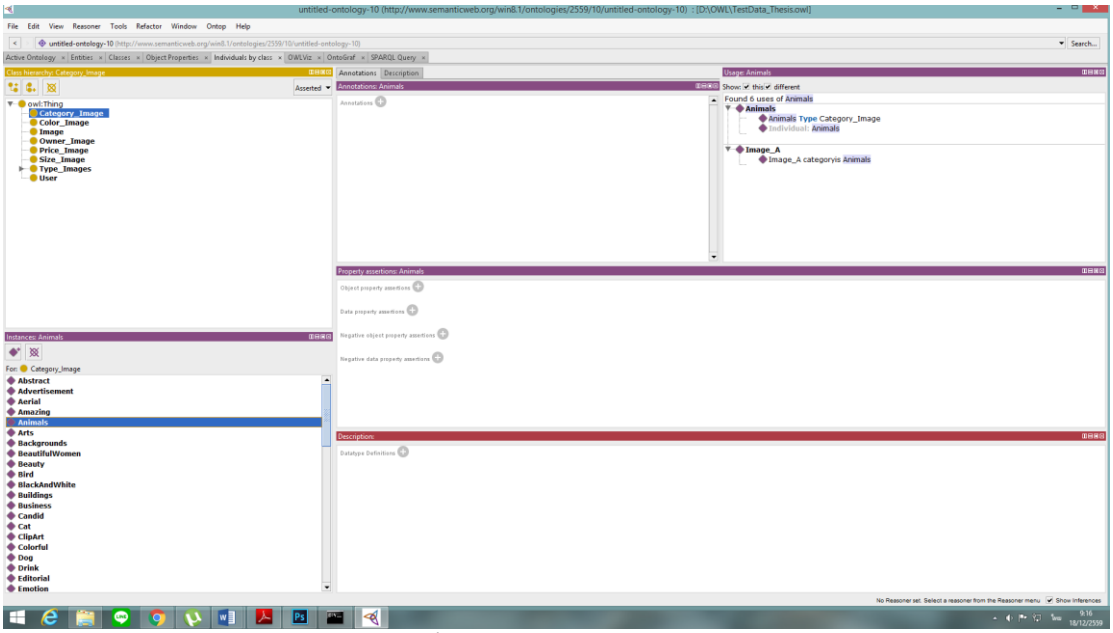
- Price
    - 10USD
    - FREE
- Size
    - 400x300
    - 640x480
    - 800x600
    - 1024x768
    - 1600x1200
    - 2272x1704
    - Exactly
    - Icon
    - LangerThan
    - Lange
    - Medium
- Type
    - Illustrations
    - Photographys
    - Vectors
- User

เมื่อเปิด OWL Project ขึ้นมา ที่แท็บคลาส จะปรากฏคลาสเริ่มต้นที่ชื่อ owl:Thing ที่ Hierarchy Tree view และเมื่อมีการเพิ่มคลาสเข้าไป ทุกคลาสจะเป็นซับคลาส ของ owl:Thing ดังภาพประกอบที่ 3 มีการเพิ่ม Class Natural เข้าไป โดยการกดปุ่ม สร้างซับคลาสและใส่ชื่อเข้าไป



ภาพประกอบที่ ผ-2 การสร้างคลาสและซับคลาส

หลังจากนั้นสร้างคลาสและซับคลาส ตามที่ได้มีการออกแบบเอาไว้ ผลลัพธ์ที่ได้ แสดงดังภาพประกอบที่ ก-3



ภาพประกอบที่ ผ-3 ลำดับชั้นของคลาสที่สร้างตามออนโทโลยีที่ออกแบบไว้

ภาคผนวก ข

ใบตอบรับการนำเสนอบทความในการประชุมวิชาการ

SPU
SRIPATUM UNIVERSITY

BANGKHEN
2410/2
PHAHOLYOTHIN RD.,
JATUJAK, BANGKOK
10900
TEL. 0 2579 1111
FAX. 0 2561 1721
www.spu.ac.th

CHONBURI CAMPUS
79 BANGNA-TRAD RD.,
KLONGTAMRU, MUANG,
CHONBURI 20000
TEL. 0 3874 3690-9
FAX. 0 3874 3700
www.east.spu.ac.th

KHON KAEN
182/12 MOO 4,
SRICHAN RD.,
NAIMUANG DISTRICT,
AMPHUR MUANG,
KHON KAEN 40000
TEL. 0 4322 4111
FAX. 0 4322 4119
www.khonkaen.spu.ac.th

ที่ มศป. 0402/3235

13 ธันวาคม 2559

เรื่อง ตอบรับการนำเสนอบทความในการประชุมวิชาการ

เรียน คุณสุชานัน อ้นถาวร

ตามที่ท่านได้ส่งบทความวิทยานิพนธ์ เรื่อง "การพัฒนาออนโทโลยีสำหรับธุรกิจขายภาพออนไลน์" เพื่อนำเสนอในงานประชุมวิชาการระดับชาติและนานาชาติ มหาวิทยาลัยศรีปทุม ครั้งที่ 11 ประจำปี 2559 เรื่อง "ผลงานวิจัยและนวัตกรรมสู่การพัฒนาที่ยั่งยืน" (National and International Sripatum Conference 2016) ในวันพุธที่ 21 ธันวาคม 2559 เวลา 8.30–16.30 น. ณ ห้อง Auditorium 1 ชั้น 14 อาคาร 40 ปีศรีปทุม มหาวิทยาลัยศรีปทุม บางเขน นั้น

ผู้ทรงคุณวุฒิ (Peer reviewers) และคณะกรรมการพิจารณาผลงานพิจารณาบทความเรื่องดังกล่าวแล้ว มีมติเห็นชอบ ให้นำเสนอบทความในการประชุมวิชาการฯ ตามวัน เวลา และสถานที่ดังกล่าวข้างต้น และจะตีพิมพ์ในรายงานสืบเนื่องการประชุมวิชาการ(Proceedings) ในรูปแบบของ CD-ROM ต่อไป

จึงเรียนมาเพื่อโปรดทราบ

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.สุบิน ยุระรัช)
ประธานคณะกรรมการพิจารณาผลงาน
การประชุมวิชาการระดับชาติและนานาชาติ
มหาวิทยาลัยศรีปทุม ครั้งที่ 11 ปีการศึกษา 2559

ฝ่ายเลขานุการคณะกรรมการพิจารณาผลงาน SPU Conference 2016
ศูนย์ส่งเสริมและพัฒนางานวิจัย มหาวิทยาลัยศรีปทุม
โทรศัพท์ 0 2579 1111 ต่อ 1331,1155, 1252
โทรสาร 0 2579 1111 ต่อ 2187
ไปรษณีย์อิเล็กทรอนิกส์ spucon2016@gmail.com

ประวัติผู้ศึกษา

ชื่อ-สกุล	นางสาวสุชานัน อ้นถาวร
วัน เดือน ปีเกิด	6 ตุลาคม 2535
สถานที่เกิด	จังหวัดชัยภูมิ
วุฒิการศึกษา	พ.ศ. 2557 เทคโนโลยีสารสนเทศและการสื่อสาร สาขาเทคโนโลยีสารสนเทศเพื่อการออกแบบ มหาวิทยาลัยศิลปากร
ตำแหน่งหน้าที่ปัจจุบัน	ตำแหน่งกราฟิกดีไซด์เนอร์
สถานที่อยู่ปัจจุบัน	55/1 ม.10 ถนน ชัยภูมิ-ตาดโตน ตำบลนาฝาย อ.เมือง จังหวัดชัยภูมิ 36000