รหัสโครงการ SUT2-204-55-12-05

รายงานการวิจัย

การเข้าถึงความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย (Access to Agricultural Knowledge by Semantic Web Technologies)

ได้รับทุนอุดหนุนการวิจัยจาก
มหาวิทยาลัยเทคโนโลยีสุรนารี

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

รหัสโครงการ SUT2-204-55-12-05

รายงานการวิจัย

การเข้าถึงความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย (Access to Agricultural Knowledge by Semantic Web Technologies)

คณะผู้วิจัย

หัวหน้าโครงการ

อาจารย์ ดร.ศุภกฤษฎิ์ นิวัฒนากูล

สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม

มหาวิทยาลัยเทคโนโลยีสุรนารี

ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ประจำงบประมาณ พ.ศ. 2555

มหาวิทยาลัยเทคโนโลยีสุรนารี

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

กันยายน 2556

กิตติกรรมประกาศ

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ประจำปีงบประมาณ พ.ศ. 2555 และงานวิจัยฉบับนี้สำเร็จลุล่วงได้ด้วยดี ด้วยเพราะได้รับการสนับสนุนอันดีจากสาขาวิชาเทคโนโลยีสารสนเทศ สถานวิจัย สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี ที่อำนวยความสะดวกด้านการดำเนินการต่าง ๆ ที่เกี่ยวข้องกับงานวิจัย ขอขอบคุณผู้ช่วยนักวิจัย คุณเจษฎา สิงห์ทองชัย คุณเอกชัย แน่นอุดร รวมถึงผู้เชี่ยวชาญ ผู้ทดสอบการใช้งานระบบที่มีส่วนช่วยในขั้นตอนต่าง ๆ ของการพัฒนาและประเมินระบบ นอกจากนี้ ยังได้รับความกรุณาอย่างดียิ่งจากผู้ทรงคุณวุฒิที่ได้เสียสละเวลาในการตรวจข้อเสนอโครงการวิจัย รวมถึงรายงานการวิจัยฉบับสมบูรณ์ และในโอกาสนี้ผู้วิจัยขอขอบคุณทุก ๆ ท่านที่มิได้ทั้งหมด ที่มีส่วนช่วยเหลือ สนับสนุน และให้กำลังใจจนกระทั่งงานวิจัยนี้ได้สำเร็จลุล่วงไปได้ด้วยดี

คุณค่าและประโยชน์อันจะพึงมีจากงานวิจัยฉบับนี้ขอมอบแด่บิดามารดา ครูอาจารย์ ครอบครัว และเพื่อนร่วมงานในสำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี ทุกท่านที่ให้กำลังใจและช่วยเหลือตลอดมา

ศุภกฤษฎ์ นิวัฒนากูล
ผู้วิจัย

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทคัดย่อ

การสืบค้นข้อมูลสารสนเทศในปัจจุบันสามารถทำได้สะดวก และ รวดเร็ว ด้วยการใช้บริการของเสิร์ชเอนจิน (Search Engine) โดยผู้ใช้สามารถระบุเงื่อนไขตามที่ต้องการ ได้แก่ คำสำคัญ (Keyword) หรือคำตามหมวดหมู่ เพื่อให้ได้ผลตามที่กำหนด เนื่องจากข้อมูลสารสนเทศมีการเปลี่ยนแปลงตลอดเวลา ดังนั้นเสิร์ชเอนจินจำเป็นจะต้องปรับปรุงดรรชนีในส่วนของข้อมูลให้ทันสมัยอยู่เสมอ ที่สำคัญผลที่ได้จากการค้นคืนสารสนเทศคือเอกสารที่มีปริมาณมากจนเกินไปซึ่งทำให้ผู้สืบค้นไม่สามารถเข้าถึงข้อมูลได้ทั้งหมด ดังนั้นการสืบค้นข้อมูลจำเป็นจะต้องใช้วิธีวัดความคล้ายคลึงกันระหว่างคำสำคัญและคำที่อยู่ในดรรชนีของคำศัพท์ต่าง ๆ เพื่อทำให้ผู้สืบค้นสามารถเข้าถึงข้อมูลได้อย่างรวดเร็ว ในงานวิจัยนี้จึงได้ศึกษาและวิเคราะห์ขั้นตอนวิธีค้นหาเอกสารการเกษตรสำหรับความรู้ไม้ผลเศรษฐกิจโดยใช้ออนโทโลยีอธิบายความหมาย ความสัมพันธ์ และโครงสร้างของข้อมูลผ่านภาษาโอดับเบิ้ลยูแอล (Web Ontology language : OWL) มีการใช้เทคนิควิธีการวัดความคล้ายคลึงกันระหว่างคำด้วยสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ด สำหรับเปรียบเทียบความคล้ายคลึงระหว่างชุดของข้อมูล และนำเสนอขั้นตอนวิธีค้นหาเชิงความหมาย เพื่อนำไปใช้ในการออกแบบและพัฒนาระบบสืบค้นเชิงความหมายสำหรับเอกสารความรู้ด้านการเกษตรของไม้ผลเศรษฐกิจในประเทศไทย ผลการทดสอบระบบพบว่า ระบบสืบค้นเชิงความหมาย มีค่าความแม่นยำร้อยละ 96.84 ค่าการจำได้ร้อยละ 83.54 และอัตราการรู้จำร้อยละ 87.55 ซึ่งการประเมินความสามารถในการใช้งานระบบสืบค้นเชิงความหมายสำหรับเอกสารความรู้ด้านไม้ผลเศรษฐกิจของประเทศไทยโดยผู้เชี่ยวชาญ มีผลประเมินโดยรวมอยู่ในระดับดี

คำสำคัญ: การสืบค้นเชิงความหมาย ออนโทโลยี สัมประสิทธิ์ความคล้ายแจ็คการ์ด
ความรู้ทางการเกษตร

Abstract

Presently, information searching can be accomplished easily and quickly with the use of search engine. This will allow the user to specify the required criteria with the use of keywords or word categories to retrieve the required results. Search engine needs to update the information index as information has been changed continuously. Significantly, the information retrieval results as in a document format are massive that makes all of the data inaccessible to the user. Consequently, information searching needs to deploy a similarity measurement method between keywords and words in an index of various terminologies to enable the user to access information rapidly. Therefore, this research studied and analyzed a search process of the agricultural document for economic fruits knowledge by employing ontology to describe relationship and information structure through Web Ontology Language (OWL). Jaccard Similarity Coefficient as a similarity measurement technique was utilized to compare the similarity between sets of data. Furthermore, a semantic search process was presented to be used in a design and development of a semantic search system for agricultural knowledge document of economic fruits in Thailand. The experimental results illustrated that this semantic search system gained 96.84 percent for precision rate, 83.54 percent for recall rate, and 87.55 percent for F-measure rate. The competency evaluation of this semantic search system for agricultural knowledge document of economic fruits in Thailand by experts was in a good level.

Keywords: Semantic search, Ontology, Jaccard Similarity Coefficient, Agricultural knowledge

สารบัญ

หน้า

กิตติกรรมประกาศ ..... ก

บทคัดย่อภาษาไทย ..... ข

บทคัดย่อภาษาอังกฤษ ..... ค

สารบัญ ..... ง

สารบัญตาราง ..... ฉ

สารบัญภาพ ..... ช

บทที่ 1 บทนำ 1

1.1 ความสำคัญและที่มาของปัญหาการวิจัย 1

1.2 วัตถุประสงค์ของการวิจัย 4

1.3 สมมติฐานการวิจัย 5

1.4 ขอบเขตของการวิจัย 5

1.5 ข้อตกลงเบื้องต้น 5

1.6 ประโยชน์ที่ได้รับจากการวิจัย 5

1.7 คำอธิบายศัพท์ 6

บทที่ 2 ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง 7

2.1 ความรู้ไม้ผลเศรษฐกิจไทย 7

2.2 การจัดเก็บและค้นคืนความรู้ 8

2.3 การค้นคืนข้อมูลด้วยเสิร์จเอนจิน 12

2.4 การสอบถามด้วยคำสำคัญ 13

2.5 เทคโนโลยีเว็บเชิงความหมาย 16

2.6 การจัดการฐานความรู้ด้วยตัวแบบออนโทโลยี 31

2.7 ภาษาการสอบถาม 37

มหาวิทยาลัยเทคโนโลยีสุรนารี

2.8 การหาเส้นทางที่สั้นที่สุด 42
2.9 งานวิจัยที่เกี่ยวข้อง....45
2.10 กรอบแนวคิดในการวิจัย....53
บทที่ 3 วิธีการดำเนินการวิจัย 57
3.1 การศึกษาความต้องการ 57
3.2 การวิเคราะห์และออกแบบระบบ 58
3.3 การพัฒนาระบบ 64
3.4 การประเมินผลระบบ 64
บทที่ 4 ผลการวิจัย 68
4.1 ออนโทโลยีไม้ผลเศรษฐกิจ....68
4.2 การวัดความคล้ายคลึงของคำค้น....73
4.3 การค้นหาเชิงความหมาย 78
4.4 การแสดงผลลัพธ์....84
4.5 ผลการประเมินระบบ 88
4.6 อภิปรายผลการวิจัย 95
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ....97
5.1 สรุปผลการวิจัย 97
5.2 ข้อเสนอแนะ 100
บรรณานุกรม....101
ภาคผนวก....108
ภาคผนวก ก ข้อมูลไม้ผลเศรษฐกิจไทย 109
ภาคผนวก ข การติดตั้ง RDF API for PHP 131
ภาคผนวก ค การหาคุณภาพและตัวอย่างแบบประเมินความสามารถของระบบ....137
ประวัติผู้วิจัย....143

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญตาราง

ตารางที่ 2.1 ตัวกระทำทางบูลีน 15
ตารางที่ 2.2 ข้อมูลสมาชิกภายในคลาส 28
ตารางที่ 2.3 ประเภทของคุณสมบัติ 28
ตารางที่ 2.4 โครงสร้างของคุณสมบัติในรูปแบบต่าง ๆ 28
ตารางที่ 2.5 ชนิดข้อมูลที่อ้างอิงตาม XML Schema datatypes 29
ตารางที่ 2.6 ประเภทข้อกำหนดของคุณสมบัติ 30
ตารางที่ 2.7 อธิบายคลาสที่พบในออนโทโลยีรถยนต์ 35
ตารางที่ 2.8 อธิบายตัวอย่างคลาส และพร๊อพเพอร์ตี้ของออนโทโลยีรถยนต์ 36
ตารางที่ 2.9 แสดงผลการสืบค้น 41
ตารางที่ 2.10 กรอบแนวคิดในการวิจัย 53
ตารางที่ 3.1 แสดงรายการโครงสร้างข้อมูลของฐานความรู้ไม้ผลเศรษฐกิจ 59
ตารางที่ 3.2 แสดงตัวอย่างคำที่ปรากฏในดัชนี และคำสืบค้น (คำถูกตามไวยากรณ์ คำสะกดผิด 62
ตารางที่ 4.1 รายละเอียดของคลาสในฐานความรู้ไม้ผลเศรษฐกิจ 68
ตารางที่ 4.2 รายละเอียดของออฟเจคพร๊อพเพอร์ตี้ในฐานความรู้ไม้ผลเศรษฐกิจ 70
ตารางที่ 4.3 รายละเอียดของดาต้าไท้ป์พร๊อพเพอร์ตี้ในฐานความรู้ไม้ผลเศรษฐกิจ 70
ตารางที่ 4.4 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลกับคำที่ถูกต้องตามไวยากรณ์ 74
ตารางที่ 4.5 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “ไม้ผลเศรษกิจ” 75
ตารางที่ 4.6 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “ทุเรียน” 75
ตารางที่ 4.7 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่ที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “มังคุด” 76
ตารางที่ 4.8 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่ที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “ลำไย” 76
ตารางที่ 4.9 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่ที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า ”ส้มเขียวหวาน” 77
ตารางที่ 4.10 แสดงการวัดประสิทธิภาพค่าสัมประสิทธิ์ความคล้ายตั้งแต่ 0.5 77
ตารางที่ 4.11 แสดงตัวอย่างการคำนวณหาค่าน้ำหนักของคำค้น 81
ตารางที่ 4.12 แสดงผลลัพธ์การคำนวณหาเส้นทางสั้นที่สุดตามโครงสร้าง 83
ตารางที่ 4.13 แสดงตัวอย่างการคำนวณหาน้ำหนักของเอกสารต่อคำค้น 84

ตารางที่ 4.14 แสดงตัวอย่างการคำนวณหาคะแนนความคล้ายเชิงความหมายของเอกสาร 85
ตารางที่ 4.15 ตัวอย่างแสดงผลการคำนวณหาคะแนนความคล้ายเชิงความหมายของเอกสาร.....86
ตารางที่ 4.16 แสดงผลเอกสารที่ถูกเลือกระหว่างวิธีการสืบค้นแบบจับคู่กับระบุโดยผู้เชี่ยวชาญ.....86
ตารางที่ 4.17 แสดงผลเอกสารที่ถูกเลือกระหว่างการสืบค้นเชิงความหมายกับระบุโดยผู้เชี่ยวชาญ 87
ตารางที่ 4.18 แสดงการวัดประสิทธิภาพผลการเปรียบเทียบระหว่างการสืบค้นแบบจับคู่กับ
การสืบค้นเชิงความหมาย 88
ตารางที่ 4.19 ข้อมูลทั่วไปของผู้ตอบแบบประเมินความสามารถระบบสืบค้น.....91
ตารางที่ 4.20 ความคิดเห็นเกี่ยวกับความสามารถของระบบสืบค้นโดยรวมในแต่ละด้าน.....92
ตารางที่ 4.21 ความคิดเห็นเกี่ยวกับด้านการประเมินความสามารถของระบบ 92
ตารางที่ 4.22 ความคิดเห็นเกี่ยวกับด้านการประเมินความถูกต้องของระบบ.....93
ตารางที่ 4.23 ความคิดเห็นเกี่ยวกับ ด้านการประเมินความง่ายในการใช้งาน 93
ตารางที่ 4.24 ความคิดเห็นเกี่ยวกับด้านการประเมินประสิทธิภาพของระบบ.....94
ตารางที่ 4.25 ความคิดเห็นเกี่ยวกับด้านการรักษาความปลอดภัย 95
ตารางที่ ก.1 ปริมาณของลำไย ปริมาณของผงกำมะถันที่ใช้และระยะเวลาที่ใช้รมลำไยสด 117
ตารางที่ ค.1 แสดงค่าดัชนีความสอดคล้อง (IOC) แบบประเมินการทำงานของระบบของผู้เชี่ยวชาญ
ที่มีต่อระบบสืบค้นไม้ผลเศรษฐกิจ 142

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญภาพ

ภาพที่ 2.1 ตัวอย่างอรรถาภิธานของ “customer” 10
ภาพที่ 2.2 ตัวอย่างข่ายงานเชิงความหมายของ “customer” 11
ภาพที่ 2.3 ภาษาและสถาปัตยกรรมของเว็บเชิงความหมาย 17
ภาพที่ 2.4 ส่วนประกอบของ Triple 21
ภาพที่ 2.5 แบบจำลองโครงสร้างข้อมูลของภาษา RDF 21
ภาพที่ 2.6 การอธิบายโครงสร้างของเมทาดาทา 22
ภาพที่ 2.7 การอธิบายความสัมพันธ์ของ Resource ในรูปแบบของอาร์ดีเอฟเอส 23
ภาพที่ 2.8 ตัวอย่างตารางข้อมูลของระบบลงทะเบียน 26
ภาพที่ 2.9 ประเภทออนโทโลยี 33
ภาพที่ 2.10 Upper Ontologyของสินค้า 34
ภาพที่ 2.11 Local Ontology ของรถยนต์ 35
ภาพที่ 2.12 ตัวอย่างข้อมูลเค้าร่างอธิบายข้อมูลรายละเอียดคลาสรถยนต์ 36
ภาพที่ 2.13 ตัวอย่างข้อมูลอินสแตนส์อธิบายรายละเอียดรถยนต์ 37
ภาพที่ 3.1 แสดงขั้นตอนการดำเนินการวิจัย 57
ภาพที่ 3.2 ระบบการสืบค้นความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย 58
ภาพที่ 3.3 ผังงานของการวัดความคล้ายคลึงของคำค้น 61
ภาพที่ 3.4 แสดงขั้นตอนการสืบค้นเชิงความหมาย 63
ภาพที่ 3.5 แสดงขั้นตอนส่วนแสดงผลลัพธ์การสืบค้น 63
ภาพที่ 4.1 โครงสร้างฐานความรู้ไม้ผลเศรษฐกิจโดยใช้ตัวแบบออนโทโลยี 69
ภาพที่ 4.2 แสดงการนิยามศัพท์ผลไม้เศรษฐกิจในเอสเคโอเอส 72
ภาพที่ 4.3 โค้ดโปรแกรมการคำนวณสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดในภาษาโปรล็อก 73
ภาพที่ 4.4 ตัวอย่างแสดงการประมวลผลโปรแกรม 74
ภาพที่ 4.5 แผนภูมิแสดงการวัดประสิทธิภาพค่าสัมประสิทธิ์ความคล้ายตั้งแต่ 0.5 78
ภาพที่ 4.6 แผนภูมิสรุปผลการประเมินความถูกต้องของการสืบค้น 88
ภาพที่ 4.7 แสดงหน้าต่างหลักของเว็บไซต์การสืบค้นเชิงความหมาย 89
ภาพที่ 4.8 แสดงตัวอย่างผลลัพธ์จากการสืบค้นเชิงความหมายกรณี 1 คำค้น 89
ภาพที่ 4.9 แสดงตัวอย่างผลลัพธ์จากการสืบค้นเชิงความหมาย 2 และ 3 คำค้น 90
ภาพที่ ก.1 ลักษณะช่อดอกลำไย 111
ภาพที่ ก.2 วงจรชีวิตหนอนคืบ 116
ภาพที่ ก.3 แมลงค่อมทอง 116

ภาพที่ ก.4 อาการใบม้วนหงิกบนช่อใบอ่อนและดอกลำไยพันธุ์ดอ 116
ภาพที่ ก.5 พันธุ์ทุเรียน.....119

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 1 บทนำ

1.1 ความสำคัญและที่มาของปัญหาการวิจัย

การเข้าถึงความรู้ที่อยู่ในระบบเครือข่ายอินเทอร์เน็ตในปัจจุบันจะใช้บริการของเสิร์ชเอนจิน (Search Engine) เช่น กูเกิ้ล (Google) ยาฮู (Yahoo) และ บิง (Bing) เป็นต้น ซึ่งเป็นบริการที่ให้ผู้ใช้สามารถเข้าถึงความรู้ที่จัดเก็บอยู่ในรูปแบบอิเล็กทรอนิกส์บนเครือข่ายอินเทอร์เน็ต แต่เนื่องจากปริมาณข้อมูล สารสนเทศ และความรู้มีปริมาณเพิ่มมากขึ้น การจัดการกับข้อมูลที่มีปริมาณมากจึงเป็นความท้าทายอย่างมากในการจัดการในการเข้าใช้ความรู้เหล่านั้น รวมถึงการค้นหาเพื่อให้ได้ผลลัพธ์ที่เหมาะสมสำหรับและตรงตามความต้องการของผู้ใช้ (Oufaida and Omar, 2009) นักวิจัยจำนวนมากจึงให้ความสนใจในการศึกษาวิจัยเกี่ยวกับระบบสืบค้นข้อมูลสารสนเทศ โดยมุ่งเน้นศึกษาหาขั้นตอนเทคนิควิธีในการสืบค้นข้อมูลให้มีความถูกต้อง รวดเร็ว และมีความเหมาะสมสามารถตอบสนองต่อความต้องการของผู้ใช้ให้มากที่สุด

ระบบสืบค้นข้อมูลสารสนเทศที่ผ่านมาในอดีตได้ใช้กลยุทธ์การค้นหาแบบตรงตามหัวข้อ โดยมีการใช้คำสำคัญเข้าไปค้นหาเพื่อให้พบข้อมูล สารสนเทศ หรือความรู้ที่เกี่ยวข้องกับคำถามของผู้ใช้ให้มากที่สุด (Christopher and Garrison,1998) และการใช้คำหลักในฐานข้อมูล (Keyword-Base) สำหรับการค้นหาเพียงอย่างเดียว ทำให้การพิจารณาถึงความหมายของเนื้อหา (Contents) ที่ได้มีประสิทธิภาพไม่เพียงพอ (Hahn and Mani, 2000) ผลลัพธ์ที่ได้จากการค้นหามีปริมาณข้อมูลหรือสารสนเทศมากเกินความจำเป็น ทำให้ยากต่อการตัดสินใจเลือกคำตอบของผู้ใช้ และเสียเวลาในการเลือกสารสนเทศหรือเอกสารที่เกี่ยวข้อง (อรวรรณ อุไรเรืองพันธุ์ และสมจิตร อาจอินทร์, 2552) ขั้นตอนและเทคนิควิธีการสืบค้นส่วนใหญ่ยังขาดความแม่นยำสำหรับการค้นหา (Craswell, Hawking and Robertson, 2001) แม้บริการเสิร์ชเอนจิน (Search Engine) จะมีการพัฒนาขั้นตอนและเทคนิควิธีสำหรับการสืบค้นให้มีความแม่นยำอย่างต่อเนื่อง แต่เนื่องจากจำนวนทรัพยากรสารสนเทศเพิ่มขึ้นอย่างรวดเร็ว และมีปริมาณมาก ทำให้การสืบค้นและการเข้าถึงข้อมูลสารสนเทศที่ตรงกับความต้องการของผู้ใช้ทำได้ยากขึ้น

บริการเสิร์ชเอนจินมีหน้าที่ในการจัดการสารสนเทศ โดยการรวบรวมข้อมูลของสารสนเทศที่กระจัดกระจายอยู่ในแหล่งต่าง ๆ ไว้เป็นกลุ่ม ๆ เพื่อให้การเข้าถึงเป็นไปได้โดยง่าย (ปิยะวัฒน์ ทองแก้ว และสมชาย ปราการเจริญ, 2553) โดยหลักการทำงานของของเสิร์ชเอนจินจะใช้คำสำคัญ เข้าไปสืบค้นสารสนเทศหรือความรู้ที่มีคำค้นอยู่ภายในเอกสารซึ่งจัดเก็บอยู่ในไฟล์ประเภทต่าง ๆ โดยการค้นหาข้อมูลจะดูว่ามีคำสำคัญที่ปรากฏอยู่ในเอกสารตรงกับข้อคำถามที่ต้องการค้นหาหรือไม่ ถ้าไม่มีระบบจะถือว่าเอกสารนั้นไม่เกี่ยวข้องหรือไม่พบเอกสาร แต่ความเป็นจริงเอกสารบางรายการอาจจะมี

ความเกี่ยวข้องกับคำร้องขอหรือคำถามที่ผู้ใช้ต้องการในเชิงความหมายถึงแม้ว่าจะไม่มีคำสำคัญปรากฏอยู่ในเอกสารที่ต้องการ ความสัมพันธ์ในลักษณะนี้เรียกว่า “ความสัมพันธ์เชิงความหมาย” ซึ่งการค้นหาข้อมูลจะพิจารณาจากแนวคิด (Concept) หรือความหมายของข้อคำถาม (Query) โดยใช้ออนโทโลยี (Ontology) ที่เป็นโครงสร้างความรู้เชิงความหมายมาแก้ปัญหาสำหรับอธิบายแนวคิดในโดเมนหนึ่ง ๆ เพื่อให้ทุกคนในโดเมนนั้นสามารถเข้าใจในความหมายของสิ่งเหล่านั้นในทางเดียวกัน การใช้ออนโทโลยีสามารถเพิ่มความครบถ้วนในการสืบค้นข้อมูลมากกว่าวิธีแบบดั้งเดิม เพราะออนโทโลยีมีคุณสมบัติในการจัดเก็บคำค้นที่มีความหมายเดียวกันหรือคล้ายคลึงกันไว้เป็นกลุ่มเพื่อใช้ขยายคำค้นจากที่ผู้ใช้ต้องการ ทำให้สามารถดึงข้อมูลที่เกี่ยวข้องกับเรื่องที่ผู้สืบค้นต้องการค้นหาได้ครอบคลุมมากขึ้นจากการเชื่อมโยงความรู้ในขอบเขตที่เกี่ยวข้องกับการสืบค้น (โรสริน อัคนิจ และคณะ, 2554)

ออนโทโลยี (Ontology) สามารถนำมาใช้ในการอธิบายความหมายของสิ่งต่าง ๆ ในขอบเขตความสนใจเรื่องใดเรื่องหนึ่ง ซึ่งสามารถเชื่อมโยงข้อมูลหรือความรู้โดยการใช้ระบบการจัดความรู้ให้เป็นระบบ (Knowledge Organization System) ซึ่งเป็นเสมือนเครื่องมือที่มีคุณลักษณะเป็นดรรชนีที่ช่วยในการจัดการและเข้าถึงเนื้อหา (Subject Access) เป็นคำศัพท์ควบคุม (Controlled Vocabularies) ที่ประกอบด้วยคำศัพท์ (Vocabularies /Terms) และไวยากรณ์ (Syntax) หรือ กฎ (Rules) เพื่อประโยชน์ในการกำหนดรูปแบบคำศัพท์ที่เหมาะสม ทำให้สามารถเข้าถึงสาระสำคัญของความรู้ที่เกี่ยวข้องได้ (Broughton, 2006; Niwattanakul, 2008)

ออนโทโลยีมีบทบาทสำคัญในการบรรยายเชิงความหมายเพื่อความเข้าใจร่วมกันในขอบเขตความรู้เรื่องใดเรื่องหนึ่งให้สอดคล้องตรงกัน โดยใช้แนวคิดเดียว (Concept) เพื่อลดหรือการตัดทอนแนวคิดหรือคำศัพท์ที่สับสน (Uschold and Gruninger, 1996) และสนับสนุนการแลกเปลี่ยน การค้นคืนข้อมูล สารสนเทศ (Hunter, 2001; Fensel, 2004) และมีความสำคัญต่อการพัฒนาระบบฐานความรู้ (Knowledge Based Systems) สำหรับการค้นหาข้อมูลที่มีความสอดคล้องและแม่นยำต่อความต้องการของผู้ใช้งาน ข้อมูลต่าง ๆ ในระบบจะถูกอธิบายด้วยภาษาเชิงความหมายที่เอื้อต่อการประมวลผลหรือค้นหาอย่างอัตโนมัติ ซึ่งประกอบไปด้วย ภาษาอาร์ดีเอฟเอส (RDF Schema :RDFS) และ ภาษาโอดับเบิ้ลยูแอล (Web Ontology Language : OWL) (W3C, www, 2006) ส่วนภาษาสำหรับการค้นหาข้อมูล เช่น สปาเก้ิล (SPARQL) และ อาร์ดีคิวแอล (RDF Data Query Language :RDQL) เป็นต้น (Prud’hommeaux and Seaborne, 2007)

การพัฒนาระบบสืบค้นหรือการค้นคืนสารสนเทศ (Information Retrieval) ได้รับความสนใจเป็นอย่างมาก สำหรับการพัฒนาระบบให้มีประสิทธิภาพ โดยเฉพาะระบบสืบค้นเกี่ยวกับความรู้และเอกสารทางการเกษตร ซึ่งใช้ฐานความรู้รูปแบบออนโทโลยีเข้ามาช่วยในการวิเคราะห์ความหมายและความสัมพันธ์ของคำศัพท์ที่เกี่ยวข้อง โดยมีการพัฒนาฐานข้อมูลอะโกรวอค (AGROVOC) ที่เป็นอรรถาภิธาน (Thesaurus) ซึ่งพัฒนาโดยองค์การอาหารและเกษตรแห่งสหประชาชาติ (FAO) และกลุ่ม

Commission of the European Communities ในปี ค.ศ. 1980 (FAO, www, 2013) ซึ่งประเทศไทยได้มีอรรถาภิธานศัพท์เกษตรไทย เป็นอรรถาภิธานด้านการเกษตรในรูปอิเล็กทรอนิกส์ ซึ่งรวบรวมคำศัพท์ด้านการเกษตรและสาขาที่เกี่ยวข้อง เช่น สิ่งแวดล้อม ป่าไม้ และประมง เป็นต้น โดยคำศัพท์แต่ละคำมีการจัดความสัมพันธ์ในลักษณะลำดับชั้น (Hierarchical Relation) เพื่อใช้ในการอธิบายเอกสารในภาษาที่เอื้อต่อการประมวลผลหรือการสืบค้น (อารีย์ ธัญกิจจานุกิจ, www, 2556)

ประเทศไทยเป็นประเทศที่ตั้งอยู่บนพื้นฐานแห่งความเป็นสังคมเกษตร ดังจะเห็นได้จากการที่ผลผลิตและรายได้จากการเกษตรยังครองความสำคัญอันดับสูงสุด เป็นสัดส่วนถึงร้อยละ 25 ของผลผลิตรวมทั้งประเทศ การจ้างแรงงานมีถึง 15.6 ล้านคน หรือประมาณร้อยละ 70 ของแรงงานทั้งประเทศ และการส่งผลิตผลการเกษตรออกไปจำหน่ายต่างประเทศทำรายได้ถึงร้อยละ 60 ของรายได้จากการส่งออกของไทยทั้งหมดในปัจจุบัน (สำนักงานเศรษฐกิจการเกษตร, 2554) อย่างไรก็ตาม การขยายตัวของการเกษตรไทยที่ผ่านมาอาศัยปัจจัยหลักที่เป็นแรงผลักดันทำให้การเกษตรเจริญเติบโตขึ้นมาโดยลำดับ คือ การเกษตรของไทยได้มีการกระจายการผลิตไปสู่พืชและผลไม้เศรษฐกิจใหม่ๆ เพิ่มขึ้นหลายชนิด เช่น ข้าว ทุเรียน และมังคุด เป็นต้น ซึ่งในแต่ละปีมีสินค้าทางการเกษตรที่ส่งออกจำหน่ายไปต่างประเทศมากกว่า 7 แสนล้านบาท โดยเฉพาะผลไม้เศรษฐกิจ ซึ่งจากรายงานข้อมูลพื้นฐานเศรษฐกิจการเกษตร ปี 2555 ของสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ มีผลไม้เศรษฐกิจที่มีการส่งออกไปจำหน่ายยังต่างประเทศสามลำดับแรก ได้แก่ ลำไย ทุเรียน และมังคุด โดยมีแนวโน้มการส่งออกที่เพิ่มขึ้นเรื่อย ๆ โดยเฉพาะการขยายการตลาดไปยังที่ประเทศจีน

โดยสรุปแล้วการพัฒนาด้านการเกษตรของไทยในระยะที่ผ่านมาปรากฏว่าผลผลิตด้านการเกษตรได้ขยายตัวในอัตราที่สูงตลอดมา และมีน้อยประเทศที่จะเทียบได้ นอกจากนั้นการเติบโตของสาขาเกษตรเป็นปัจจัยสำคัญต่อการเพิ่มพูนฐานะและรายได้ให้แก่เกษตรกรและบุคคลที่เกี่ยวข้อง และมีส่วนสำคัญต่อการแก้ปัญหาความยากจนของประชากรทั้งประเทศ แต่ในปัจจุบันความเหลื่อมล้ำในฐานะรายได้ระหว่างภาคเกษตรและภาคเศรษฐกิจอื่นๆ และระหว่างในภูมิภาคส่วนต่าง ๆ ของประเทศยังมีปัญหาอยู่อีกมากโดยเฉพาะการเข้าถึงการศึกษาหรือความรู้เฉพาะทางด้านการเกษตรโดยมีเทคโนโลยีสารสนเทศมาเป็นปัจจัยสำคัญที่จะมาช่วยสำหรับการเพิ่มผลผลิต ลดต้นทุน หรือการแก้ไขปัญหาจากการผลิตให้มีความสอดคล้องกับความต้องการต่อเศรษฐกิจการเกษตรของประเทศที่มีการเจริญเติบโตอย่างต่อเนื่อง ฉะนั้นผู้ผลิตหรือเกษตรกรจำเป็นอย่างยิ่งที่จะต้องมีความรู้ เพื่อที่จะใช้ความรู้ในการแก้ไขปัญหาได้ทันท่วงที ถึงแม้ในปัจจุบันจะมีระบบอินเตอร์เน็ตที่ใช้เสิร์จเอนจินให้บริการการค้นหาเอกสารหรือความรู้ที่มีความสะดวกและรวดเร็ว แต่เนื่องจากระบบเสิร์จเอนจินที่มีอยู่ในปัจจุบันไม่ได้ออกแบบมาเพื่อรองรับการค้นหาเอกสารหรือความรู้ทางการเกษตรเป็นการเฉพาะเจาะจง และระบบการค้นหายังเป็นการสืบค้นจากคำสำคัญเป็นหลักโดยไม่ได้พิจารณาจากความหมายของคำสำคัญ จึงทำให้ผลลัพธ์ที่ได้ไม่ตรงกับความต้องการ หรือยังมีเอกสารที่เกี่ยวข้องอยู่ในฐานข้อมูลหรือฐานความรู้ที่ไม่ได้ถูกค้นหาดึงมาใช้ประโยชน์

จากความต้องการของเกษตรกรและผู้เกี่ยวข้องที่ต้องการเข้าถึงความรู้ทางการเกษตรที่อยู่กระจัดกระจายบนอินเตอร์เน็ตโดยใช้บริการของเสิร์ชเอ็นจิน แต่ปัญหาสำคัญของการใช้เสิร์ชเอนจินในปัจจุบันได้แก่จำนวนเอกสารความรู้ที่คืนคืนมาได้มีปริมาณมากเกินไปจะกระทั่งไม่สามารถจะเข้าถึงได้ทั้งหมด การค้นหาแบบใช้คำหลัก (Keywords) ตามที่ผู้ใช้ระบุทำให้เอกสารที่เกี่ยวข้องอีกจำนวนมากที่อยู่ในฐานความรู้อื่น ไม่ได้ถูกดึงขึ้นมาเนื่องจากไม่มีคำค้นตามที่ผู้ใช้ระบุ ซึ่งเป็นปัญหาของคำที่เขียนต่างกันแต่มีความหมายเดียวกัน นอกจากนี้การจำแนก การจัดทำดัชนี การกำหนดความสัมพันธ์ระหว่างความรู้ต่าง ๆ มีความสำคัญต่อการค้นหาเพื่อให้มีความแม่นยำ ตรงตามความต้องการของผู้ใช้มากขึ้น จึงเป็นเรื่องที่สำคัญที่จะต้องนำมาประยุกต์ใช้สำหรับการจัดารเอกสารความรู้ทางการเกษตรที่อยู่กระจัดกระจายบนเครือข่ายอินเตอร์เน็ต

จากความสำคัญและปัญหาของการสืบค้นสารสนเทศที่ผู้วิจัยกล่าวมาข้างต้นจะพบว่าการวิจัยและพัฒนาระบบสืบค้นสารสนเทศได้มีการรวมเอาเทคนิคหลากหลายด้านมาประกอบกัน เพื่อให้การสืบค้นมีประสิทธิภาพและตอบสนองต่อความต้องการของผู้ใช้มากที่สุด โดยใช้เทคโนโลยีที่พัฒนาขึ้นที่สามารถรองรับการสืบค้นเอกสารโดยการสืบค้นเชิงความหมาย ซึ่งผ่านการสร้างองค์ความรู้แบบออนโทโลยี (Ontology) ดังนั้นผู้วิจัยจึงมีความสนใจที่จะทำการที่จะนำเสนอขั้นตอนวิธีการค้นหาเชิงความหมายสำหรับเอกสารความรู้การเกษตรของผลไม้เศรษฐกิจในประเทศไทย โดยใช้ออนโทโลยีอธิบายความหมาย ความสัมพันธ์และโครงสร้างของข้อมูลผ่านภาษาโอดับเบิ้ลยูแอล (Web Ontology language : OWL) ซึ่งมุ่งเน้นการนำเสนอเทคนิควิธีการสกัดข้อมูลด้วยการวัดความคล้ายคลึงกันของแจ็คการ์ด (Jaccard Similarity Coefficient) และขั้นตอนวิธีการสืบค้นเชิงความหมายโดยใช้เทคนิควิธีการหาระยะทางที่สั้นที่สุดจากโครงสร้างความสัมพันธ์เชิงความหมายของคำค้น โดยทำให้ผลการค้นหาเอกสารความรู้ด้านการเกษตรมีความถูกต้อง แม่นยำตรงกับความต้องการของผู้ใช้ อันจะมีประโยชน์สำหรับเกษตรกรไทยและผู้ที่สนใจที่จะก้าวไปแข็งขันสู่โลกยุคสังคมเศรษฐกิจฐานความรู้ และจะเป็นฟันเฟืองอันหนึ่งที่จะขับเคลื่อนการเกษตรของไทยไปสู่การเกษตรแบบยั่งยืนเพื่อการแข่งขัน และเศรษฐกิจพอเพียงอย่างสมดุล

1.2 วัตถุประสงค์ของการวิจัย

1) เพื่อออกแบบและพัฒนาฐานความรู้การเกษตรด้านไม้ผลเศรษฐกิจโดยใช้ตัวแบบออนโทโลยี
2) เพื่อออกแบบขั้นตอนวิธีการสืบค้นเชิงความหมายสำหรับเอกสารความรู้การเกษตร
3) เพื่อพัฒนาระบบสืบค้นเชิงความหมายสำหรับฐานความรู้การเกษตร

1.3 สมมติฐานการวิจัย

1) ผลการประเมินระบบสืบค้นเชิงความหมายสำหรับเอกสารความรู้ด้านไม้ผลเศษฐกิจของประเทศไทย มีค่าความแม่นยำ (Precision) ค่าการจำได้ (Recall) และอัตราการรู้จำ (F-measure) ไม่น้อยกว่าร้อยละ 80
2) การประเมินความสามารถในการใช้งานระบบสืบค้นเชิงความหมายสำหรับเอกสารความรู้ด้านไม้ผลเศรษฐกิจของประเทศไทยโดยผู้เชี่ยวชาญ มีผลประเมินในภาพรวมระดับดีขึ้นไป

1.4 ขอบเขตของการวิจัย

ในการวิจัยนี้ได้กำหนดขอบเขตการวิจัยเฉพาะไม้ผลเศรษฐกิจของประเทศไทยจากรายงานข้อมูลพื้นฐานเศรษฐกิจการเกษตร ปี 2552 ในสามลำดับแรก ได้แก่ ลำไย ทุเรียน และมังคุด โดยกำหนดให้มีรายการความรู้ที่จัดเก็บไม่น้อยกว่า 600 รายการ

1.5 ข้อตกลงเบื้องต้น

ในงานวิจัยนี้ผู้วิจัยได้กำหนดข้อตกลงเบื้องต้นของการวิจัย ดังนี้

1) เอกสารความรู้ไม้ผลเศรษฐกิจที่นำมาเป็นตัวอย่างจะค้นหาจากระบบเครือข่ายอินเตอร์เน็ตด้วยเสิร์ชเอ็นจิน แล้วเลือกรายการความรู้ที่มีการเผยแพร่ตั้งแต่ปี พ.ศ. 2550 เป็นต้นมา โดยเลือกรายการที่มีคำอธิบายครบถ้วนตามความต้องการ
2) ฐานความรู้ไม้ผลเศรษฐกิจที่พัฒนาขึ้น จะจัดเก็บเฉพาะคำอธิบายเอกสารความรู้ และลิงก์เชื่อมโยงไปยังเอกสารความรู้เท่านั้น
3) ระบบการสืบค้นความรู้ทางการเกษตรด้านไม้ผลเศรษฐกิจจะจัดทำส่วนต่อประสานผู้ใช้เฉพาะส่วนการสืบค้นเอกสารความรู้ในฐานความรู้ที่พัฒนาขึ้นมาเท่านั้น โดยผลลัพธ์ที่แสดงผ่านเว็บเบราว์เซอร์ ที่ประกอบด้วย ชื่อเอกสารความรู้ ชื่อผู้เขียน ลิงก์เชื่อมโยงไปยังเอกสารรายละเอียดโดยย่อ และจัดลำดับแสดงรายการผลลัพธ์

1.6 ประโยชน์ที่ได้รับจากการวิจัย

1) ได้ฐานความรู้การเกษตรสำหรับไม้ผลเศรษฐกิจของประเทศไทย ซึ่งสามารถเชื่อมโยงไปยังความรู้เกี่ยวกับไม้เศรษฐกิจไม่น้อยกว่า 600 รายการ
2) ได้ขั้นตอนวิธีการสืบค้นเชิงความหมายสำหรับความรู้การเกษตรของไม้ผลเศรษฐกิจที่สนับสนุนภาษาไทย

3) ได้ระบบสืบค้นเชิงความหมายสำหรับความรู้การเกษตรของไม้ผลเศรษฐกิจ เพื่อให้นักวิชาการเกษตร เกษตรกร และผู้สนใจ ทำการสืบค้นผ่านเครือข่ายอินเทอร์เน็ต

1.7 คำอธิบายศัพท์

1) การสืบค้นเชิงความหมาย (Semantic Search) หมายถึง กระบวนการค้นหาที่มีการนำคำสำคัญ (Keyword) ไปหาคำสำคัญอื่น ๆ ที่เกี่ยวข้อง แล้วใช้กลุ่มคำสำคัญเหล่านั้นในการสืบค้นความรู้ที่มีการจัดทำอธิบายในฐานความรู้

2) ความรู้ไม้ผลเศรษฐกิจ (Fruit Knowledge) หมายถึง องค์ความรู้ที่เกี่ยวข้องกับไม้ผลเศรษฐกิจเฉพาะ ลำไย ทุเรียน และมังคุด ซึ่งจัดเก็บอยู่ในไฟล์ประเภทต่าง ๆ เช่น เวิร์ด (Word) พีดีเอฟ (PDF) และไฟล์ภาพ เป็นต้น

3) ออนโทโลยีความรู้การเกษตร (Ontology Agricultural Knowledge) หมายถึง การอธิบายความหมายของสิ่งต่าง ๆ และสามารถจัดหมวดหมู่เอกสารของความรู้ในขอบเขตของไม้ผลเศรษฐกิจของประเทศไทย

4) คลังคำ (Word Repository) หมายถึง การนำกลุ่มคำมาจัดความสัมพันธ์ระหว่างคำเหล่านั้น เพื่อใช้ประโยชน์ในการช่วยการค้นคืนความรู้

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 2 ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้เน้นการเข้าถึงความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย ที่สามารถบรรยายกลุ่มข้อมูลของไม้ผลเศรษฐกิจของประเทศไทย โดยผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องตามลำดับ ดังนี้

2.1 ความรู้ไม้ผลเศรษฐกิจไทย

2.2 การจัดเก็บและค้นคืนความรู้

2.3 การค้นหาข้อมูลด้วยเสิร์ชเอ็นจิน

2.4 การสอบถามด้วยคำสำคัญ

2.5 เทคโนโลยีเว็บเชิงความหมาย

2.6 การจัดการฐานความรู้ด้วยตัวแบบออนโทโลยี

2.7 ภาษาการสอบถามข้อมูล

2.8 การหาเส้นทางที่สั้นที่สุด

2.9 งานวิจัยที่เกี่ยวข้อง

2.10 กรอบแนวคิดการวิจัย

2.1 ความรู้ไม้ผลเศรษฐกิจไทย

สำหรับความรู้ไม้ผลเศรษฐกิจไทยมีอยู่กระจัดจายโดยทั่วไปทั้งในรูปแบบเอกสารและรูปแบบข้อมูลอิเล็กทรอนิกส์ ซึ่งจากการศึกษาจะพบโครงสร้างของความรู้ทางการเกษตรประกอบไปด้วยข้อมูลที่สำคัญ ดังนี้ (พงษ์ศักดิ์ อังกสิทธิ์ และคณะ, 2542; เดโชดม ภัทรศัย, 2543;)

- **ลักษณะทางพฤกษศาสตร์**

ข้อมูลลักษณะทางพฤกษศาสตร์ประกอบไปด้วยชื่อภาษาไทย ชื่อภาษาอังกฤษ ชื่อวิทยาศาสตร์ วงศ์ นอกจากนี้แล้วยังได้มีการอธิบายเพิ่มเติมถึงลักษณะต้น ใบ ดอก ผล และเมล็ด ซึ่งจะมีความแตกต่างกันในไม้ผลแต่ละชนิด

- **การปลูกและการดูแลรักษา**

สำหรับการปลูกและดูแลรักษาจะประกอบด้วยข้อมูลพันธุ์ที่ใช้ในการปลูก การเตรียมพื้นที่ปลูก รวมถึงการบอกถึงคุณลักษณะทางสภาพแวดล้อมที่เหมาะสมสำหรับพืชแต่ละชนิด ในการปลูกจะมี

คำแนะนำระยะการปลูกที่เหมาะสม หลังจากปลูกแล้วก็จะมีการกล่าวถึงการดูแลรักษาทั้งเรื่องของการใส่ปุ๋ยในช่วงระยะเวลาต่าง ๆ เช่น ก่อนออกดอก ระหว่างออกดอก ช่วงติดผล เป็นต้น การตัดแต่งกิ่งเพื่อช่วยในการติดดอกออกผลที่ดีขึ้น เพื่อช่วยในการป้องกันโรคและแมลง

- **โรคและแมลงศัตรู**

สำหรับข้อมูลที่สำคัญประกอบด้วย ชื่อโรคและแมลง อาการที่พบจากการทำลายของโรคและแมลง วิธีการป้องกันโรคและแมลง ซึ่งข้อมูลโรคและแมลงศัตรูพืชอาจจะมีส่วนเกี่ยวข้องกับสารเคมีที่ใช้ในการป้องกับและกำจัดศัตรูพืช ซึ่งข้อมูลสารเคมีส่วนใหญ่จะประกอบด้วยชื่อสารออกฤทธิ์ ชื่อสารออกฤทธิ์ทางการค้า รายละเอียดการใช้สารเคมีทางการค้า การแก้พิษทางเคมี และวิธีการเก็บสารเคมี

- **การเก็บเกี่ยวและการเก็บรักษา**

ข้อมูลการเก็บเกี่ยวจะประกอบด้วยช่วงระยะเวลาที่เหมาะสมสำหรับการเก็บเกี่ยว วิธีการเก็บเกี่ยว และการปฏิบัติหลังการเก็บเกี่ยว ซึ่งอาจมีการให้ข้อมูลเกี่ยวกับการแปรรูป ได้แก่ วิธีการแปรรูป กระบวนการแปรรูป และส่วนประกอบที่นำมาแปรรูป

- **การตลาด**

สำหรับข้อมูลสำหรับการตลาดจะเกี่ยวข้องแหล่งรับซื้อ ราคา ช่องการจำหน่าย รวมถึงข้อมูลการส่งออกขายยังต่างประเทศ

ในการศึกษาความรู้ทางการเกษตร ผู้วิจัยได้ศึกษาข้อมูลไม้ผลเศรษฐกิจของประเทศไทยที่มีการส่งออกไปจำหน่ายยังต่างประเทศ 3 ลำดับแรก ได้แก่ ลำไย ทุเรียน มังคุด จากการรายงานข้อมูลพื้นฐานเศรษฐกิจการเกษตร ของสำนักงานเศรษฐกิจการเกษตร ปี 2552 โดยทำการศึกษาข้อมูลที่เกี่ยวกับลักษณะทางพฤกษศาสตร์ การปลูก การปฏิบัติดูแลรักษา โรคและแมลงศัตรูสำคัญ และการจัดการหลังการเก็บ และการวิจัยด้านไม้ผลเศรษฐกิจของไทย เพื่อนำข้อมูลดังกล่าวนี้ไปใช้ในการออกแบบฐานความไม้ผลเศรษฐกิจของไทย สำหรับรายละเอียดความรู้ของไม้ผลเศรษฐกิจไทยที่สำคัญ ได้แก่ ลำไย ทุเรียน และมังคุด ปรากฏตามภาคผนวก ก

2.2 การจัดเก็บและค้นคืนความรู้

ระบบจัดเก็บและค้นคืนความรู้ (Knowledge Storage and Retrieval System) เป็นกระบวนการในการรวบรวมรายละเอียดของความรู้ที่จัดเก็บในฐานความรู้ เพื่อให้สามารถค้นความรู้กลับคืนมาได้อย่างสะดวกและรวดเร็ว และตรงกับความต้องการของผู้ใช้มากที่สุด ดังนั้น การจัดเก็บและค้นคืนความรู้จึงมีความสำคัญอย่างยิ่งต่อผู้ที่ต้องการค้นหาความรู้ที่ต้องการ เนื่องจากว่าผู้ใช้ต้องการเข้าถึงความรู้ได้อย่างสะดวก รวดเร็ว และตลอดเวลา ผู้ใช้มีความต้องการในการได้รับความรู้ที่ถูกต้องและตรงกับความต้องการเพื่อนำไปใช้ประโยชน์ได้ทันต่อเหตุการณ์ นอกจากนี้ ยังต้องสนับสนุน

ให้ผู้ใช้สามารถประเมิน แยกแยะ ทำความเข้าใจ เชื่อมโยงความรู้ เพื่อการศึกษา แสวงหา และติดตามความรู้ต่าง ๆ ได้ด้วยตนเอง

เพื่อให้การจัดเก็บและค้นคืนความรู้ได้อย่างสะดวก รวดเร็ว และทันต่อความต้องการ จึงต้องมีการจัดทำความรู้ให้เป็นระบบ ซึ่งระบบการจัดทำความรู้ให้เป็นระบบ (Knowledge Organization System : KOS) ตามที่ฮอดจ์ (Hodge, 2000) ได้นำเสนอไว้ มี 3 ประเภท ดังนี้

- **รายการข้อความ (Term lists)**

รายการข้อความ หมายถึง การรวบรวมคำสำคัญ หรือคำหลักในขอบเขตทั่วไปหรือเฉพาะเจาะจงไว้ในรูปแบบแฟ้มข้อมูล หรืออาจจะเป็นฐานข้อมูล เพื่อที่จะค้นหาและนำไปใช้ได้โดยสะดวก ซึ่งรายการข้อความ อาจประกอบด้วย

1) รายการหลักฐาน (Authority files) เป็นการแสดงรายการของคำหรือข้อความที่ใช้ในการควบคุมชื่อที่แตกต่างกันสำหรับโดเมนใดโดเมนหนึ่ง เช่น รายชื่อสัตว์น้ำ รายชื่ออุปกรณ์ประกอบอาหาร เป็นต้น

2) อภิธานศัพท์ (Glossaries) เป็นรายการของคำหรือข้อความพร้อมกับความหมายของคำหรือข้อความเหล่านั้น

3) พจนานุกรม (Dictionaries) เป็นรายการคำตามลำดับตัวอักษรพร้อมกับความหมาย และอาจมีรายละเอียดเพิ่มเติม เช่น คำเหมือน คำตรงข้าม ประเภทของคำ เป็นต้น

4) อักขรานุกรมภูมิศาสตร์ (Gazetteers) เป็นรายการของชื่อสถานที่ ซึ่งอาจเป็นชื่อเมือง ภูเขา ถนน ธนาคาร สถานีบริการน้ำมันเชื้อเพลิง เป็นการรวมรวมรายชื่อสถานที่ไว้ เพื่อให้สะดวกต่อการค้นหาในภายหลัง

- **การจัดกลุ่มและจำแนก (Classifications and Categories)**

การจัดกลุ่มและจำแนก เป็นการจัดมาตรฐานของการจำแนกสิ่งต่าง ๆ ที่อยู่ในโลกนี้ เพื่อให้ผู้ที่เกี่ยวข้องนำไปใช้ในการควบคุมและจัดหมวดหมู่ของวัตถุต่าง ๆ อันจะทำให้มีความเป็นมาตรฐานเดียวกัน จะทำให้ค้นหาได้สะดวกยิ่งขึ้น

การจัดกลุ่มและจำแนก อาจประกอบด้วย

1) หัวเรื่อง (Subject headings) เป็นกลุ่มของข้อความควบคุมเพื่อเป็นตัวแทนหัวเรื่องของรายการหรือวัตถุต่าง ๆ ที่ได้รวบรวมไว้ ดังเช่น ระบบหัวเรื่องของห้องสมุดรัฐสภาอเมริกัน หรือที่รู้จักในชื่อ LCSH (Library of Congress Subject Headings) เพื่อใช้ในการควบคุมหัวเรื่องของรายการหนังสือในห้องสมุด

2) รูปแบบการจัดกลุ่ม, อนุกรมวิธาน และรูปแบบการจำแนก (Classification schemes, Taxonomies, and Categorization schemes) เป็นการจัดกลุ่มและจำแนกตามหัวเรื่องที่ได้กำหนดไว้ตามลำดับชั้น การจัดกลุ่มของวัตถุตามคุณลักษณะเฉพาะ ดังเช่น อนุกรมวิธานสัตว์บก จะเป็นระบบการจำแนกประเภทของสัตว์บกตามลำดับชั้นที่ได้กำหนดไว้

- **รายการความสัมพันธ์ (Relation lists)**

รายการความสัมพันธ์ เป็นระบบการจัดทำความรู้ให้เป็นระบบที่ซับซ้อนมากขึ้นกว่าสองกลุ่มที่ได้กล่าวมาแล้วข้างต้น เนื่องจากว่ารายการข้อความเป็นเพียงคำหรือข้อความเท่านั้น ไม่มีการสร้างความสัมพันธ์ระหว่างคำหรือข้อความเหล่านั้น ในขณะที่การจัดกลุ่มและการจำแนกจะมีการสร้างความสัมพันธ์ระหว่างคำหรือข้อความเหล่านั้นแบบเป็นลำดับชั้น แต่รายการความสัมพันธ์จะสามารถสร้างความสัมพันธ์ระหว่างคำหรือข้อความมากกว่านั้น ทำให้การจัดการรายการความสัมพันธ์จะมีความยุ่งยากและซับซ้อนมากกว่า

รายการความสัมพันธ์ที่ใช้กันโดยทั่วไปในปัจจุบัน ประกอบด้วย

1) อรรถาภิธาน (Thesaurus) เป็นการกำหนดความสัมพันธ์ระหว่างคำต่าง ๆ ที่ได้รวบรวมไว้ อาจเป็นการจัดเก็บคำหรือข้อความในขอบเขตใดขอบเขตหนึ่ง ซึ่งความสัมพันธ์ที่สามารถแสดงระหว่างคำหรือข้อความเหล่านั้น อาจมีข้อมูลเกี่ยวกับคำเหมือน (Synomym) คำตรงข้าม (Antonym) คำพ้องรูปพ้องเสียง (Homonym) คำที่ใช้ทั่วไป (Generic term) คำที่เป็นส่วนหนึ่งของ (Part/whole term) คำที่เกี่ยวข้อง (Asssociative term) และข้อมูลอื่น ๆ จากตัวอย่างอรรถาภิธานของคำว่า “customer” จะทำให้ทราบถึงข้อมูลต่าง ๆ ที่เกี่ยวข้องกับคำนี้ ดังเช่น จะได้ทราบว่าคำเหมือน คำตรงข้าม มีคำใดบ้าง ตัวอย่างดังภาพที่ 2.1

Main Entry:	**customer** [kuhs-tuh-mer] Show IPA
Part of Speech:	*noun*
Definition:	buyer of goods, services
Synonyms:	client, clientele, consumer, habitué, patron, prospect, purchaser, regular shopper
Antonyms:	owner

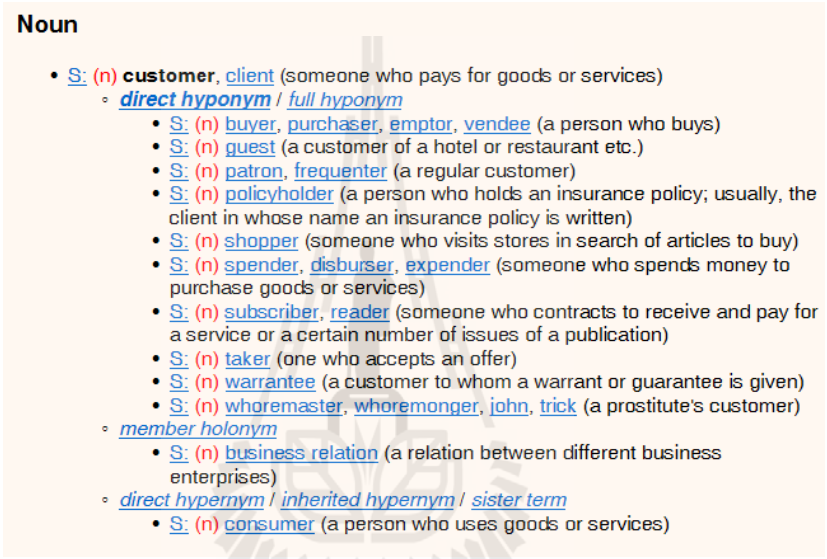
ที่มา : www.thesaurus.com

ภาพที่ 2.1 ตัวอย่างอรรถาภิธานของ “customer”

ในการจัดการความรู้ทางเกษตร ได้มีการกำหนดคำศัพท์เพื่อช่วยในการค้นหาความรู้ทางเกษตรในรูปแบบอรรถาภิธานศัพท์สากล AGORVOC Thesaurus ที่พัฒนาโดยองค์การอาหารและเกษตรแห่งสหประชาติ (Food and Agricultural of United Nations) ซึ่งรองรับ 22 ภาษา รวมทั้งภาษาไทย

(FAO, www, 2013) สำหรับประเทศไทยแล้ว ได้มีการพัฒนาอรรถาภิธานศัพท์เกษตรไทยเพื่อเป็นคำศัพท์มาตรฐานสำหรับกำหนดดรรชนีสารสนเทศเกษตรของไทย โดยจะช่วยเพิ่มประสิทธิภาพในการสืบค้นสารสนเทศทางเกษตรในการขยายขอบเขตคำค้น ค้นคำศัพท์ที่เป็นคำพ้องเสียงและคำพ้องความหมายซึ่งปัจจุบันมีคำศัพท์ที่ให้บริการมากกว่า 28,000 คำ (อารีย์ ธัญกิจจานุกิจ, www, 2556)

2) ข่ายงานความหมาย (Semantic network) เป็นการจัดเก็บข้อมูลในลักษณะกราฟซึ่งประกอบด้วยโหนด (Node) ซึ่งเป็นตัวแทนแนวคิดหรือสิ่งใดสิ่งหนึ่ง และเส้นเชื่อม (Edge) ซึ่งใช้แทนความสัมพันธ์เชิงความหมายระหว่างโหนดหรือแนวคิดนั้น ๆ ตัวอย่างดังภาพที่ 2.2



Noun

- S: (n) **customer**, client (someone who pays for goods or services)
  - ***direct hyponym*** / *full hyponym*
    - S: (n) buyer, purchaser, emptor, vendee (a person who buys)
    - S: (n) guest (a customer of a hotel or restaurant etc.)
    - S: (n) patron, frequenter (a regular customer)
    - S: (n) policyholder (a person who holds an insurance policy; usually, the client in whose name an insurance policy is written)
    - S: (n) shopper (someone who visits stores in search of articles to buy)
    - S: (n) spender, disburser, expender (someone who spends money to purchase goods or services)
    - S: (n) subscriber, reader (someone who contracts to receive and pay for a service or a certain number of issues of a publication)
    - S: (n) taker (one who accepts an offer)
    - S: (n) warrantee (a customer to whom a warrant or guarantee is given)
    - S: (n) whoremaster, whoremonger, john, trick (a prostitute's customer)
  - *member holonym*
    - S: (n) business relation (a relation between different business enterprises)
  - *direct hypernym* / *inherited hypernym* / *sister term*
    - S: (n) consumer (a person who uses goods or services)

(ที่มา http://wordnetweb.princeton.edu/)

ภาพที่ 2.2 ตัวอย่างข่ายงานเชิงความหมายของ “customer”

3) ออนโทโลยี (Ontology) ในทางวิทยาการคอมพิวเตอร์ สตูเดอร์ เบนจามินส์ และเฟนเซล (R. Studer, V.R. Benjamins, and D. Fensel, 1998) ได้ให้ความหมายไว้ว่า ออนโทโลยี คือ รูปแบบข้อกำหนดที่ชัดเจนของแนวคิดที่มีการแบ่งปันหรือใช้ร่วมกัน ซึ่งจากความหมายดังกล่าวออนโทโลยีจะต้องมี รูปแบบ (Formal) ที่เครื่องจักรสามารถประมวลผลได้ ความชัดเจน (Explicit) ของรูปแบบที่มีการกำหนดไว้ และ แบ่งปัน (Shared) ตามที่ได้มีการตกลงที่จะใช้งานร่วมกัน

ในงานวิจัยนี้ได้นำแนวทางของอรรถาภิธานศัพท์ไปใช้ในการจัดทำคลังคำเพื่อช่วยในการค้นคืนข้อมูลความรู้ไม้ผลเศรษฐกิจไทย โดยใช้รายการคำจาก AGROVOC Thesaurus ขององค์การอาหารและเกษตรแห่งสหประชาชาติ (FAO) ซึ่งเป็นอรรถาภิธานศัพท์เกษตรที่เป็นมาตรฐานในระดับสากล เป็นคลังคำศัพท์ตั้งต้นเฉพาะที่เป็นคำศัพท์ภาษาต่างประเทศ และ อรรถาภิธานศัพท์เกษตรไทย ซึ่งเป็น

อรรถาภิธาน (Thesaurus) ด้านการเกษตรในรูปแบบอิเล็กทรอนิกส์ที่มีการรวบรวมคำศัพท์ด้านการเกษตรและสาขาที่เกี่ยวข้องกับการเกษตร

2.3 การค้นคืนข้อมูลด้วยเสิร์จเอนจิน

การค้นคืนข้อมูลบนระบบเครือข่ายอินเตอร์เน็ตในปัจจุบันจะใช้ซอฟต์แวร์ที่เรียกว่าเสิร์จเอ็นจิน (Search Engine) ซึ่งเป็นระบบซอฟต์แวร์ที่มีซอฟต์แวร์หุ่นยนต์ (Robot Software) ซึ่งบางครั้งเรียกว่า Spider หรือ Crawler ทำหน้าที่ท่องไปในเว็บไซต์ต่าง ๆ ในอินเทอร์เน็ตเพื่อรวบรวมเอกสารบนเว็บ โดยนำมาสร้างเป็นฐานดัชนีสำหรับการสืบค้นเอกสารบนเว็บ เพื่อทำให้การสืบค้นทำได้อย่างรวดเร็ว ดังนั้นกระบวนการจัดเตรียมดัชนี จึงเป็นขั้นตอนสำคัญอย่างยิ่งสำหรับการสืบค้นข้อมูลบนเว็บ

ศุภชัย ตั้งวงศ์ศานต์ (2553 : 545-549) ได้อธิบายไว้ว่า เสิร์จเอ็นจิน คือ จักรกลการค้นหา หรือระบบการสืบค้นเอกสารที่สร้างขึ้นเพื่อช่วยการค้นหาสารสนเทศ ซึ่งจัดเก็บอยู่ในระบบคอมพิวเตอร์ที่เป็นของส่วนตัว องค์กร หรือบนเครือข่าย หรือบนเครือข่ายอินเทอร์เน็ต เพื่อให้ตรงกับบทของการบรรยาย เสิร์จเอ็นจินที่จะสามารถสืบค้นได้เร็ว นอกจากต้องมีระบบคอมพิวเตอร์ที่มีประสิทธิภาพสูงแล้วที่สำคัญเสิร์จเอ็นจินได้มีดรรชนีของคำศัพท์ต่าง ๆ ที่บ่งบอกตำแหน่งทำให้การค้นหาได้รวดเร็ว แต่เนื่องจากเว็บไซต์บนอินเทอร์เน็ตมักจะเปลี่ยนแปลงข้อมูลอยู่เสมอ รวมทั้งที่เกิดขึ้นใหม่และที่หายไป ทำให้เสิร์จเอ็นจินต้องปรับปรุงดรรชนี และส่วนอื่นให้ทันสมัย

บริษัทเสิร์จเอ็นจินที่สำคัญ ได้แก่ Google, Yahoo!Search, Microsoft และ Ask.com (ศุภชัย ตั้งวงศ์ศานต์, 2553) ซึ่งเป็นของบริษัทเองที่พัฒนาขึ้นมาเพื่อการสืบค้น แต่บางบริษัทก็อาจจะอ้างเป็นเสิร์จเอ็นจินเช่นกันที่จะส่งผลให้เกิดรายได้จากการโฆษณาและธุรกิจที่เกี่ยวเนื่องตามมา ซึ่งหลักการทำงานของเสิร์จเอ็นจินประกอบด้วย 3 ขั้นตอน ดังนี้ 1) การแสวงหาข้อมูล 2) การสร้างดรรชนี และ 3) การค้นหาสารสนเทศ

โดยทั่วไปผู้ใช้บริการเสิร์จเอ็นจินจะตั้งคำถามที่เป็นการใส่คำสำคัญเป็นคำ หรือเป็นชุด ส่วนการตั้งคำถามเป็นภาษาตอบโต้ หรือแม้แต่พูดใส่เข้าไปในคอมพิวเตอร์ก็จะเป็นแนวทางของการพัฒนา IR โดยใช้เทคนิคของปัญญาประดิษฐ์ (Artificial Intelligence) มาช่วย ซึ่งในกรณีของการใส่คำสำคัญสามารถใส่ได้หลายรูปแบบ เช่น คำสำคัญทั่วไป คำสำคัญจำเพาะ คำสำคัญที่เป็นตัวเลข การใช้คำเชื่อม AND, OR, NOT และการสืบค้นด้วยคำใกล้เคียง เป็นต้น

ทุกวันนี้ นอกจาก Search Engine เช่น Google, Yahoo ที่เป็นการสืบค้นทั่วไปบนอินเทอร์เน็ตที่ให้บริการฟรีแล้ว ทำให้เป็นที่นิยมของคนทั่วไปถึงขนาดเป็นที่กล่าวกันว่าหากมีปัญหา หรือข้อสงสัยใดถาม Google! ด้วยเหตุผลที่ว่า Google จะให้คำตอบของคำสอบถามได้มากมายแล้ว ยังให้บริการโดยไม่คิดค่าใช้จ่าย อย่างไรก็ตามยังมีคลังข้อมูลที่ให้บริการแต่ไม่ฟรีอีกมากมาย เช่น การบริการ e-Journals ของ IEEE, ACM, JETP (Journal of Experimental and Theoretical Physics) เป็นต้น

(ศุภชัย ตั้งวงศ์ศานต์, 2553) เสิร์จเอ็นจินโดยส่วนมากจะมีผลประโยชน์ทางการค้ามาเกี่ยวข้อง อาจจะมาในรูปแบบการลงโฆษณา และการขายสินค้าตามมาทุกครั้งของการแสดงผล บ่อยครั้งการจัดอันดับความสำคัญจึงไม่เป็นไปตามสิ่งที่ควรจะเป็น กลายเป็นตามค่าตอบแทนที่ลูกค้าของเสิร์จเอ็นจินนั้นได้จ่ายให้ และกลายเป็นจัดอยู่หน้าต้น ๆ ของการแสดงผล

สรุปได้ว่า การแสวงหาข้อมูลบนเว็บเป็นกระบวนการแรกของเสิร์จเอ็นจิน ที่จะได้มาซึ่งข้อมูล เนื่องจากเว็บหรือเครือข่ายสื่อสารมีการเชื่อมโยงสลับซับซ้อนทั้งเป็นระเบียบ และไม่เป็นระเบียบคล้ายตาข่ายแมงมุม โดยระบบการสืบค้นเอกสารที่สร้างขึ้นเพื่อช่วยการค้นหาสารสนเทศ จะเป็นระบบการสืบค้นเอกสารที่สร้างขึ้นเพื่อช่วยการค้นหาสารสนเทศ แสดงผลลัพธ์ตามความคำสอบถามของผู้ใช้ ซึ่งงานวิจัยนี้ผู้วิจัยได้มีการออกแบพัฒนาส่วนระบบการสืบค้นเอกสาร การค้นหาเชิงความหมายขึ้นเพื่อให้เข้าถึงฐานความรู้ด้านการเกษตร

2.4 การสอบถามด้วยคำสำคัญ

การค้นหาข้อมูลในเสิร์ชเอนจินจะเป็นการสอบถามด้วยคำสำคัญ (Keyword) ซึ่งเป็นรูปแบบที่ง่ายที่สุดที่และมีผู้ใช้นิยมใช้มากที่สุด การสอบถามแต่ละครั้งอาจประกอบด้วยคำสำคัญเดี่ยว ๆ เช่น Bangkok หรือคำสำคัญหลาย ๆ คำเรียงลำดับเป็นวลี เช่น Bangkok Symphony Orchestra หรือคำสำคัญที่เชื่อมต่อกันในรูปแบบอื่น ๆ เช่นคำสอบถามเป็นภาษาธรรมชาติ การสอบถามด้วยคำสำคัญมี (ศุภชัย ตั้งวงศ์ศานต์, 2553) ดังนี้

1) คำสำคัญเป็นคำเดี่ยว

การสืบค้นลักษณะนี้ ผู้ใช้มักจะนึกถึงคำสำคัญที่เป็นคำเฉพาะในเอกสาร หรือสิ่งเป้าหมายเป็นหลัก เช่น ในการสืบค้นข้อมูลของเมืองหลวงของประเทศไทย คำสำคัญที่เป็นคำเฉพาะคงเป็น Bangkok หรือสนใจในการสืบค้นข้อมูลสัตว์ขนาดใหญ่ที่ปรากฏในโลกคำสำคัญที่เป็นคำเฉพาะคงเป็น Dinosaur แม้ว่าสัตว์ประเภทนี้จะสูญพันธ์ไปแล้วก็ตาม หรือเราอาจจะสอบถามด้วยคำลักษณะอื่นก็ได้ ระบบการสืบค้นในเป้าหมายก็ควรจะสืบค้นมาให้ได้ตามความต้องการของผู้ใช้ แต่อย่างไรก็ตามคำในภาษาอังกฤษ มักจะมีการแปลงรูปเพื่อการใช้ของประโยค เช่นการเติม s เติม ed เติม ing เช่นกรณีของ Retrieve กลายเป็น Retrieves, Retrieved, Retrieving หรือบางทีก็เปลี่ยนโฉมกลายเป็น Retrieval เมื่อใช้เป็นคำนาม อีกประการณ์หนึ่งคำในภาษาอังกฤษมีทั้งการเขียนที่เป็นตัวเล็กและตัวใหญ่ ขึ้นอยู่กับตำแหน่งและสถานการณ์ โดยทั่วไประบบการสืบค้นก็จะไม่ได้แยกแยะตัวเล็กตัวใหญ่โดยถือว่าเป็นคำในความหมายเดียวกัน เช่น gates หรือ Gates แต่หากต้องการค้นเฉพาะ Gates ปรากฏลักษณะนี้ผู้ใช้ก็ควรจะใช้สัญลักษณ์ “ ปิดหัวท้ายเพื่อเน้นเป็นการเฉพาะ เช่น “Gates” ระบบก็จะถึงเอาคำเฉพาะคำ Gates ออกมาให้ปรากฏ โดยไม่รวม gates, gate, gating ในผลลัพธ์ด้วย คำใน

ภาษาอังกฤษบางคำที่เขียนด้วยอักษรเฉพาะต่อไปนี้ หากจะให้ระบบการสืบค้นถือเป็นคำเดี่ยว จะต้องปิดหัวท้ายด้วย “

สำหรับการค้นหาคำภาษาไทยนั้น อาจจะไม่มีการต่อเติมท้าย หรือแปลงรูปอย่างภาษาอังกฤษ แต่การสืบค้นด้วยคำสำคัญเดี่ยว ๆ ก็จะมีความซับซ้อนอีกลักษณะหนึ่ง กล่าวคือ ภาษาไทยจะเขียนเป็นประโยคด้วยคำติดกันโดยไม่มีเว้นวรรค จะเว้นวรรคก็ต่อเมื่อเป็นคำวลีหรือประโยคคำสืบค้นด้วยคำเดี่ยวคงไม่ใช่เป็นเพียงหนึ่งพยางค์ แต่เป็นหลายพยางค์ในหนึ่งคำ เช่น โรงเรียน เป็นต้น

2) คำสำคัญเป็นวลี

การสืบค้นด้วยคำสำคัญที่เรียงต่อเป็นวลีเป็นรูปแบบที่ให้ความหมายและประโยชน์ต่อการสืบค้น เช่น การค้นหาด้วยเทอม Bangkok Symphony Orchestra ก็จะได้ผลลัพธ์ของวงดนตรีที่เป็นชื่อเฉพาะข้างต้น ผลลัพธ์ที่แสดงย่อมมีจำนวนน้อยกว่าการค้นหาด้วยแต่ละคำโดด ๆ อันจะเป็นการได้รายการมากแต่ประโยชน์น้อย ในการสืบค้นแบบวลีจะคำนึงถึงลำดับก่อนหลังของคำสอบถามที่ปรากฏในวลีเดียวกัน โดยจะดึงเฉพาะวลีที่ตรงกับลำดับการปรากฏของคำที่ผู้ใช้ระบุไว้เท่านั้น เช่นการค้นวลี Information Retrieval ก็จะมีเฉพาะเอกสาร หรือสิ่งเป้าหมายที่ปรากฏของคำว่า Information ตามด้วย Retrieval แต่ย่อมไม่ปรากฏในสิ่งที่เป็น Retrieval ตามด้วย Information

การเลือกคำสำคัญในการค้นด้วยวลีก็เป็นเรื่องที่ต้องพิจารณา โดยทั่วไป การเลือกใส่คำยิ่งมากผลลัพธ์ย่อมออกมายิ่งน้อย (ศุภชัย ตั้งวงศ์ศานต์, 2553) แต่ก็เป็นการเฉพาะมากขึ้น แต่อย่างไรก็ตามการที่ใส่คำสอบถามมาก ๆ ก็ไม่จำเป็นเสมอไปว่าจะได้ผลดีกว่า ทั้งนี้ขึ้นอยู่กับความหมายของคำสำคัญในการสืบค้นนั้น ๆ ตัวอย่างเช่น ต้องการสืบค้นสถานที่พักที่ดี ๆ ในกรุงเทพฯ แทนที่จะใช้คำสำคัญในการสืบค้นว่า Luxury Hotels Bangkok ก็จะได้รายการที่ต้องการและเป็นที่น่าพอใจ แต่กลับกลายเป็นว่าใช้คำยาว ๆ เช่น Really Nice Places to Spend the Night in Bangkok ก็คงจะได้ผลลัพธ์อีกรูปแบบหนึ่ง และก็ไม่เป็นที่ต้องการ เพราะระบบปัจจุบันยังไม่ฉลาดพอที่จะเข้าใจในความหมาย ของคำสอบถาม ที่เรียกว่า Semantic Search

นอกจากคำในวลีที่เขียนติดกัน หรือคั่นด้วยประเภท Stop-word แล้ว ในการสืบค้นบ่อยครั้งต้องการค้นหาคำสำคัญที่อยู่ใกล้เคียงกัน (Proximity) โดยมีคำอื่นคั่นกลางได้ตั้งแต่ 1 คำขึ้นไป และคำคั่นกลางนั้นจะรวม Stop-words หรือไม่ก็ได้ ดังตัวอย่างจะแสดงตำแหน่งของคำสอบถาม 2 คำที่สามารถอยู่ใกล้กันหรือห่างกันไม่เกิน 5 คำ เช่น การสืบค้นคำ Information Retrieval ควรจะตรงกับข้อความ คือ ....Information with the power of retrieval..... หมายความว่าให้สืบค้นเนื้อหาที่มีคำ 2 คำในประโยคเดียวกัน ซึ่งอาจจะเป็นคำในชื่อเรื่อง สาระสังเขป หรือคำสำคัญก็ได้

3) คำสำคัญด้วยตัวเชื่อมทางตรรกะ

Keyword Search เป็นการสืบค้นข้อมูลโดยอาศัยหลักการของบูลีน (Boolean Search) Keyword หรือคำเฉพาะ คือ คำ หรือ กลุ่มคำที่ใช้แทนความหมายของเนื้อเรื่อง นิยมใช้ในการทำเครื่องมือช่วยสืบค้นข้อมูลจากแหล่งต่าง ๆ หรือบทความวารสาร หรือแหล่งข้อมูลที่กว้างอย่าง www เสิร์จเอ็นจิน หลายแห่งได้มีการใช้หลักการสืบค้นแบบบูลีน โดยอาศัยตัวกระทำของบูลีน (Boolean Operation) ประกอบด้วย AND, OR และ NOT ทำหน้าที่เชื่อมคำหรือกลุ่มคำที่ต้องการค้นหา โดยเป็นผลให้จำนวนผลลัพธ์ ที่ได้จากการสืบค้นลดน้อยลง หรือเป็นไปตามที่เราต้องการมากขึ้น ตัวกระทำดังกล่าวมีการใช้ที่แตกต่างกัน ดังแสดงในตารางที่ 2.1

ตารางที่ 2.1 ตัวกระทำทางบูลีน

Boolean	เครื่องหมาย	ความหมาย
AND	+	AND Operator แทนสัญลักษณ์ในการสืบค้นด้วยเครื่องหมาย + เมื่อเชื่อมด้วย and ระบบจะทำการค้นหาระเบียนที่คำหรือข้อความทั้งสองปรากฏอยู่ด้วยกัน เช่น คำสั่ง + มหาวิทยาลัย หมายถึง ให้ค้นหาเอกสารที่มีทั้งคำว่า “คำสั่ง” และ คำว่า“มหาวิทยาลัย” โดยจะต้องมีคำทั้งสองคำนี้ปรากฏอยู่ทั้งคู่จะมีเพียง คำใดคำหนึ่งไม่ได้
OR	&	OR Operator แทนสัญลักษณ์ในการสืบค้นด้วยเครื่องหมาย & เมื่อเชื่อมด้วย or ระบบจะทำการค้นหาระเบียนที่มีคำใดคำหนึ่งหรือข้อความใดข้อความหนึ่งที่ปรากฏ เช่น คำสั่ง & มหาวิทยาลัย หมายถึงให้ค้นหาเอกสารที่คำว่า “คำสั่ง” หรือคำว่า “มหาวิทยาลัย” เพียงคำหนึ่งหรือทั้งสองคำก็ได้
NOT	-	Not Operator แทนสัญลักษณ์ในการสืบค้นด้วยเครื่องหมาย - เมื่อเชื่อมด้วย not ระบบจะทำการค้นหาระเบียนที่คำหรือข้อความแรกปรากฏอยู่โดยที่ระเบียนนั้น จะต้องไม่มีคำหรือข้อความที่สองอยู่ในเขตข้อมูลใดๆ เช่น คำสั่ง – มหาวิทยาลัย หมายถึงให้ค้นหาเอกสารที่มีคำว่า “คำสั่ง” แต่ในเอกสารนั้นต้อง ไม่มีคำว่า “มหาวิทยาลัย” ปรากฏอยู่

4) การสอบถามด้วยภาษาธรรมชาติ

ภาษาธรรมชาติ (Natural Language) หมายถึงการเขียนคำสอบถามเป็นวลี หรือประโยคที่มีความหมาย แต่คำสำคัญที่เป็นวลีและต่อด้วยคำเชื่อมทางตรรกะ AND , OR, NOT ก็เป็นประโยคที่มีความหมายในตัวเอง ตัวอย่างเช่น

“Events of the 2004 Tsunami in Indonesia”

หากตัดคำ Stop-words ทั้งหลายออกไป ก็จะเหลือคำ Events, 2004, Tsunami, Indonesia คำเหล่านี้คือคำสำคัญในการสอบถามจักรกลการค้นหา

การใช้ภาษาธรรมชาติในการสืบค้นเอกสารหรือสิ่งเป้าหมายเป็นเรื่องที่มนุษย์เราต้องการใช้ แต่ด้วยข้อจำกัดของเทคโนโลยี การจะใช้ภาษาธรรมชาติที่ทำการสืบค้นโดยตรงยังทำไม่ได้มีประสิทธิภาพมากนัก จึงมาสร้างการสืบค้นด้วยภาษาที่ไม่เป็นธรรมชาติ แทน เช่น การให้เป็นคำสำคัญตามกลไกที่สร้างตัวดรรชนี การสร้าง XQL, XQurey เป็นต้น

สรุปได้ว่า จากการศึกษาทฤษฎีการสอบถามด้วยคำสำคัญ (Keyword) ผู้วิจัยได้ทำการเลือกใช้วิธีการสอบถามด้วยคำสำคัญคำเดียว (Key word) เพราะการสืบค้นลักษณะนี้ ผู้ใช้มักจะนึกถึงคำสำคัญที่เป็นคำเฉพาะในเอกสารหรือสิ่งเป้าหมายเป็นหลัก

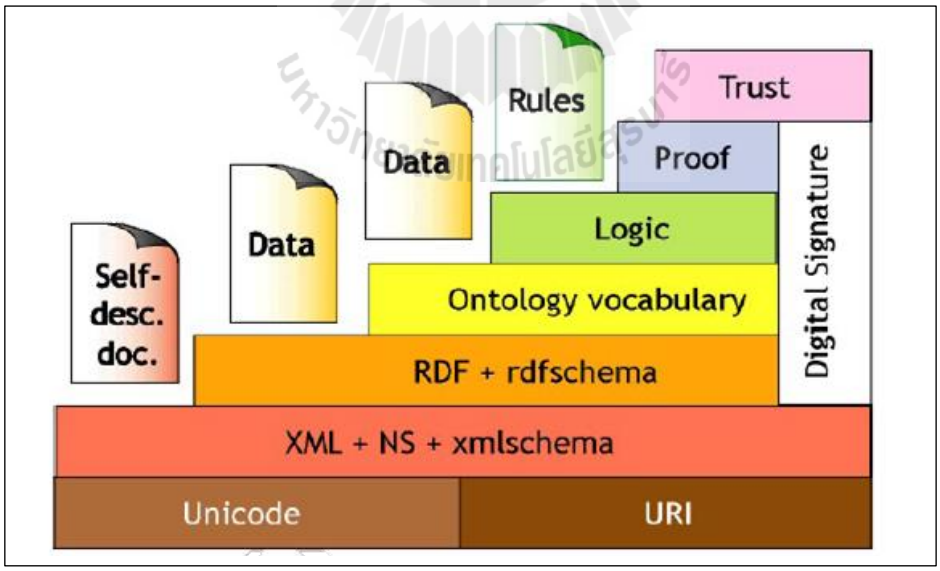
2.5 เทคโนโลยีเว็บเชิงความหมาย

เว็บเชิงความหมาย (Berners-Lee et al., 2001) เป็นวิวัฒนาการของเวิลด์ไวด์เว็บ (W3C, www, 2009) ซึ่งเกี่ยวข้องกับข้อมูล และการบริการบนเว็บไซต์ โดยสร้างความเป็นไปได้ที่เว็บไซต์ และคอมพิวเตอร์จะสามารถเข้าถึงความต้องการของผู้ใช้งานได้ เว็บแบบเชิงความหมายบรรจุไปด้วยเซ็ทของหลักการออกแบบการทำงานร่วมกัน และความหลากหลายของเทคโนโลยีพื้นฐาน บางส่วนของเว็บแบบสื่อความหมายแสดงให้เห็นถึงความเป็นไปได้ ที่จะรองรับกับเทคโนโลยี หรือสามารถนำมาใช้ได้จริงในอนาคต ซึ่งจะประกอบด้วย Resource Description Framework (RDF) ความหลากหลายของการสับเปลี่ยนข้อมูล (เช่น RDF/XML, N3, Turle, N-Triples) และเครื่องหมาย เช่น RDF Schema (RDFS) และ Web Ontology Language (OWL) ซึ่งสิ่งเหล่านี้มุ่งหมายเพื่อเตรียมการถึงส่วนประกอบของการจำกัดความ และความรู้ที่ได้รับ ฉะนั้นเว็บแบบเชิงความหมาย คือเว็บข้อมูลที่คอมพิวเตอร์สามารถเข้าใจ และประมวลผลข้อมูลได้เอย่างอัตโนมัติ หลักการที่ทำให้เกิดการพัฒนาเว็บแบบเชิงความหมายอย่างต่อเนื่องคือ สาเหตุจากการที่เว็บในปัจจุบันเป็นเพียงเว็บที่ใช้แสดงเอกสารหรือแสดงข้อมูลผ่านทางเว็บบราวเซอร์เท่านั้น (NECTEC,www, 2009) คอมพิวเตอร์ไม่เข้าใจเพียงรู้ว่าข้อมูลเหล่านั้นคืออะไร มีความสัมพันธ์กันอย่างไร แนวทางของเว็บแบบเชิงความหมายจะช่วยแก้ปัญหาดังกล่าว ก็คือการนำเสนอข้อมูลเกี่ยวกับ Thing อาทิ คน สถานที่ การเกษตร สัตว์ ฯลฯ โดยใช้รูปแบบของการอธิบายข้อมูลด้วยเมทาดาทา (Metadata) ที่เป็นมาตรฐานเดียวกันให้กับคอมพิวเตอร์ได้รับรู้

เว็บเชิงความหมาย (Vong, 2003; อติพงศ์ จุลละโพธิ, 2548) หรือเว็บที่อิงข้อมูล เป็นเว็บที่มีลักษณะที่ทำให้เครื่องคอมพิวเตอร์สามารถเข้าใจข้อมูล (Data) ร่วมกันได้ โดยการกำหนดโครงสร้างในการอธิบายข้อมูลที่เป็นมาตรฐานเดียวกันให้กับเครื่องคอมพิวเตอร์รับรู้และเข้าใจข้อมูลร่วมกัน เว็บเชิงความหมาย (มารุต บูรณรัช และ เทพชัย ทรัพย์นิธิ, 2550) มีวิธีการที่จะทำให้คอมพิวเตอร์สามารถเข้าถึงเว็บไซต์ที่สัมพันธ์กันได้โดยอัตโนมัติ และยังสามารถดึงข้อมูลที่เกี่ยวข้องกันออกมาจากฐานข้อมูลแบบเสปรดชีท (SpreadSheet) หรือไฟล์รูปภาพได้ด้วย

จากความหมายของเว็บแบบเชิงความหมายที่กล่าวมา สามารถสรุปได้ว่า เว็บแบบสื่อความหมายมีจุดประสงค์เพื่อเพิ่มขีดความสามารถให้กับเว็บไซต์ในปัจจุบัน เป็นที่รวมของเทคโนโลยีหลักในการใช้ Metadata เพื่อบ่งบอกถึงความชัดเจน ครอบคลุมถึง ออนโทโลยี เหตุผล และเว็บแบบเชิงความหมายยังมีแนวคิดพื้นฐานสำหรับหลักการสำคัญที่ทำให้เครื่องคอมพิวเตอร์สามารถจะประมวลสารสนเทศบนเว็บไซต์ได้เองอัตโนมัติ

เทคโนโลยีเว็บเชิงความหมาย เป็นเครื่องมือสำหรับบูรณาการและจัดระเบียบข้อมูลเชิงความหมาย ทำให้มีการเชื่อมโยงเครือข่ายของข้อมูลบนเว็บไซต์ต่างๆ และส่งผลให้การสืบค้นเป็นไปอย่างมีประสิทธิภาพและสามารถสร้างความสัมพันธ์ให้กับข้อมูลที่มาจากแหล่งที่ต่างกันเป็นฐานข้อมูลที่เชื่อมโยงกัน แนวคิดหลักของเทคโนโลยีเชิงความหมาย คือ การพัฒนาเทคโนโลยีที่ทำให้การสืบค้นเป็นไปอย่างชาญฉลาดด้วยเทคโนโลยีและภาษาต่างๆ ที่ใช้กับเว็บเชิงความหมาย (พนิดา ตันศิริ, 2553)



ภาพที่ 2.3 ภาษาและสถาปัตยกรรมของเว็บเชิงความหมาย

หลักการพื้นฐานของเทคโนโลยีเว็บแบบเชิงความหมายสามารถอธิบายได้จากลำดับชั้นของการออกแบบและแนวความคิดของเว็บแบบเชิงความหมาย แสดงดังภาพที่ 2.3 (Antoniou and Harmelen, 2008)

- **Extensible Markup Language (XML)**

เอ็กซ์เอ็มแอล (Extensive Markup Language: XML) เป็นภาษาที่ใช้สำหรับการเขียนภาษาที่ใช้นิยามความหมายของเอกสารหรือข้อมูล (markup) โดยที่เอกสารนั้นมีการใช้เมทาดาทา (Metadata) เพื่อบอกหน้าที่และประเภทของข้อมูลของส่วนต่างๆ ของเอกสารนั้นได้โดยชัดเจน การเพิ่มเมทาดาทาเข้าไปในเอกสารสามารถทำให้โครงสร้างของเอกสารชัดเจนขึ้น และทำให้การประมวลผลเอกสารเป็นไปโดยง่ายและไม่จำเป็นที่จะต้องอาศัยมนุษย์เพื่อตีความเอกสาร เอ็กซ์เอ็มแอลพัฒนาสำหรับการแลกเปลี่ยนข้อมูลของระบบที่แตกต่างกัน โดยมีการแยกส่วนของข้อมูลและส่วนแสดงผลออกจากกันเพื่อความยืดหยุ่นในการเลือกรูปแบบการแสดงผลข้อมูล ดังรายละเอียดลักษณะการใช้งาน โครงสร้าง และองค์ประกอบของเอกสารเอ็กซ์เอ็มแอล (W3C, www, 2004; Hunter et al, 2002; สราวุธ อ้อยศรีสกุลม, 2544)

- *ลักษณะการใช้งานของเอ็กซ์เอ็มแอล*

ลักษณะการใช้งานของเอ็กซ์เอ็มแอล 1) ใช้ในการออกแบบการพัฒนาโปรแกรมเว็บเพื่อการจัดส่งข้อมูลสารสนเทศ 2) ใช้ในการสร้างเมทาดาทา (Metadata) หรือแท็กเพื่อบอกหน้าที่และประเภทข้อมูลของส่วนต่าง ๆ ในเอกสาร ช่วยให้มีเอกสารมีโครงสร้างชัดเจนขึ้น 3) ช่วยให้การประมวลผลเอกสารเป็นไปโดยง่าย 4) เอ็กซ์เอ็มแอลใช้งานในส่วนประมวลผลข้อมูลแล้วส่งผ่านให้โปรแกรมประยุกต์ เพื่อส่งไปยังแหล่งเก็บข้อมูลต่าง ๆ ซึ่งเอ็กซ์เอ็มแอลมีความสามารถในการสนับสนุนแอพพลิเคชั่นหลายรูปแบบ และ 5) เอกสารเอ็กซ์เอ็มแอลสามารถจัดเก็บในระบบฐานข้อมูล และกำหนดโครงสร้างของเอกสารได้

- *โครงสร้างของเอกสารเอ็กซ์เอ็มแอล*

เอกสารเอ็กซ์เอ็มแอลจะแบ่งโครงสร้างออกเป็น 3 ส่วน ซึ่งประกอบด้วย

1) ส่วนหัวของเอกสาร จะเรียกว่า Prolog หรือเรียกอีกอย่างหนึ่งว่า XML Declaration หมายถึงส่วนหัวของเอกสารที่ประกาศการใช้ เป็นโครงสร้างส่วนแรกของเอกสารเอ็กซ์เอ็มแอลเรียกว่า Prolog เป็น Optional (มีหรือไม่มีก็ได้) ใน Prolog มีองค์ประกอบอีก 2 ส่วนคือ XML Declaration และ Document type Declaration

XML Declaration - เป็นการระบุ Version ของ XML Specification ซึ่งควรประกาศไว้เสมอ และยังอาจประกอบด้วย Encoding Declaration เพื่อระบุแบบแผนการเข้ารหัสตัวอักษร เช่น

UTF-8 เป็น Default Character Encoding Scheme เป็นตัวแทนของตัวอักษรส่วนใหญ่ในภาษาอังกฤษรวมถึงภาษาไทยด้วย

ตัวอย่าง:

```
<?xml version="1.0" encoding="UTF-8"?>
```

Document Type Declaration - ประกอบด้วย Markup Code ที่กำหนดกฎการเขียน หรือ Document Type Definition (DTD)

ตัวอย่าง:

<! DOCTYPE TroubleTicket SYSTEM "TroubleTicket.dtd">

- TroubleTicket หมายถึง root element
- TroubleTicket.dtd หมายถึงให้เช็คโครงสร้างจากไฟล์ TroubleTicket.dtd

ตัวอย่างที่ 1:

```
1: <?xml version="1.0" ?>
2: <! DOCTYPE TroubleTicket SYSTEM "TroubleTicket.dtd">
3: <TroubleTicket ticketID="T7426284">
4: <Description> Customer product </Description>
5: </TroubleTicket>
```

2) ส่วนพื้นที่แสดงข้อมูลในเอกสารเอ็กซ์เอ็มแอล ที่เรียกว่า Document Element ประกอบด้วยเนื้อหาแบบโครงสร้าง Hierarchical Tree หรือเป็นส่วนของ Content จริง Document Element เป็น Single Element ที่สามารถประกอบด้วย Sub Elements และ External Entities ไม่จำกัดจำนวน กล่าวอีกนัยหนึ่ง Document Element ก็คือ Root Element ของเอกสารนั่นเอง เอ็กซ์เอ็มแอลมีลักษณะเป็น Case Sensitive นั่นคือตัวพิมพ์เล็กและตัวพิมพ์ใหญ่มีความหมายแตกต่างกัน มี Tag เปิดและปิด ถ้า Empty Tag จะใช้

ในบรรทัดที่ 1 และ 2 จะเป็น Prolog ของเอกสาร ส่วนบรรทัดที่ 3 – 5 จะเป็น Elements ที่มีข้อมูลอยู่ ซึ่งข้อมูลตัวนี้สามารถเรียกดูได้

3) ส่วนของการเขียนคำอธิบายคำสั่ง หรือฟังก์ชั่นการทำงาน เรียกว่า Epilog

ตัวอย่าง:

<?!—End id =1-->

- *องค์ประกอบของเอกสารเอ็กซ์เอ็มแอล*

1) XML Element เป็นส่วนของการแสดงอิลิเมนต์ที่เกี่ยวข้องกับการใช้งาน จะประกาศเป็นค่าตัวแปร

2) Attribute เป็นส่วนที่ใช้ในการแสดงคุณสมบัติของสิ่งที่เราสนใจ

3) Entity เป็นส่วนแสดงถึงสิ่งที่เกิดขึ้น หรือมีการเปลี่ยนแปลง

4) Comment ใช้แสดงหมายเหตุของโปรแกรม จะไม่มีการตีความ

5) Processing Instructions เป็นส่วนของการแทรกข้อมูลเข้าไปยังแอพพลิเคชั่น

6) CDATA Sections เป็นส่วนของการป้องกันความผิดพลาดของการอ่านเอกสารเมื่อข้อความภายในแท็กมีสัญลักษณ์ที่เหมือนกับที่เอ็กซ์เอ็มแอลใช้งานอยู่

- **Resource Description Framework (RDF)**

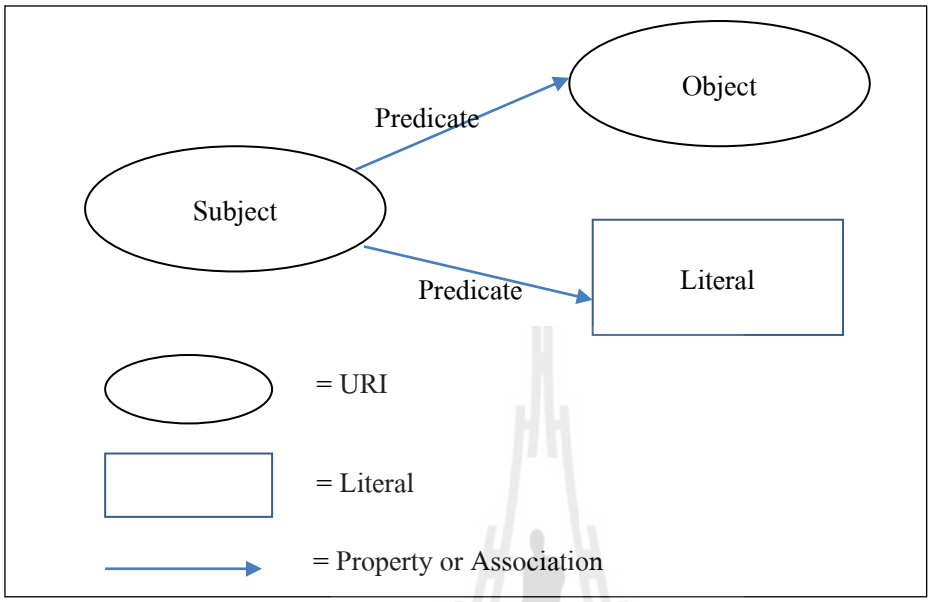
อาร์ดีเอฟ (Record Description Framework: RDF) เป็นภาษามาตรฐานที่ใช้อธิบายลักษณะและความสัมพันธ์ของข้อมูลคล้ายกับแผนภาพอีอาร์ (Entity Relationship Model: ER) ที่ใช้สำหรับอธิบายข้อมูลสารสนเทศที่จัดเก็บจากแหล่งต่าง ๆ และเป็นตัวกลางทำหน้าที่ในส่วนช่วยให้เครื่องคอมพิวเตอร์เข้าใจข้อมูลที่บรรยาย และสามารถตีความหมายข้อมูลตามที่ต้องการได้ อาร์ดีเอฟใช้โครงสร้างและรูปแบบของเอ็กซ์เอ็มแอล ซึ่งเป็นการพัฒนาต่อยอดจากเอ็กซ์เอ็มแอลหรือ XML-Base Syntax ด้วยเหตุนี้อาร์ดีเอฟจึงอยู่ชั้นบนกว่าเอ็กซ์เอ็มแอล

อาร์ดีเอฟถือได้ว่าเป็นตัวกลางที่จะช่วยให้เครื่องคอมพิวเตอร์สามารถเข้าใจข้อมูลร่วมกันได้ เครื่องคอมพิวเตอร์สามารถที่จะเข้าใจข้อมูลที่ถูกบรรยายด้วยภาษาอาร์ดีเอฟ และสามารถตีความข้อมูลตามที่ต้องการได้อาร์ดีเอฟ มีวิธีการบรรยายรายละเอียดของข้อมูลโดยประกอบด้วย 3 ส่วนที่เรียกว่า Triple ดังนี้คือ ส่วนทรัพยากร (Subject) ส่วนคุณสมบัติของทรัพยากร (Predicate) และส่วนค่าของคุณสมบัติ (Object) แสดงดังภาพที่ 2.4 (Klyne and Carroll, 2004)

1) Subject คือ ส่วนทรัพยากรหรือส่วนสิ่งที่สนใจ ใช้สัญลักษณ์แทนด้วยโหนดที่เป็นวงรี ซึ่งแทนด้วย URI

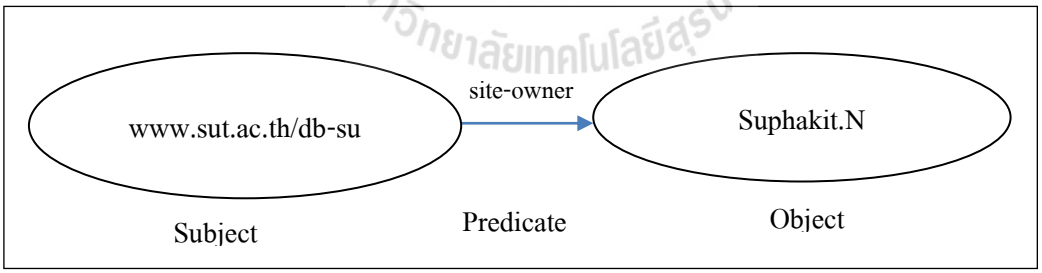
2) Predicate คือ คุณสมบัติของทรัพยากร ใช้สัญลักษณ์แทนด้วยลูกศรชี้จากส่วน Subject ชี้ไปยังส่วน Object หรือ Literal ซึ่งแทนด้วย URI

3) Object คือ ค่าของคุณสมบัติ ใช้สัญลักษณ์แทนด้วยโหนดที่เป็นวงรี ซึ่งแทนด้วย URI แต่ถ้าหากเป็นค่าของอักขระจะเรียกว่า Literal ซึ่งจะใช้สัญลักษณ์แทนด้วยโหนดที่เป็นสี่เหลี่ยม



ภาพที่ 2.4 ส่วนประกอบของ Triple

สำหรับการแทนข้อมูลตามโครงสร้างอธิบายข้อมูลของภาษาอาร์ดีเอฟ สามารถแทนข้อมูลดังภาพที่ 2.5 (W3C, www, 2004)



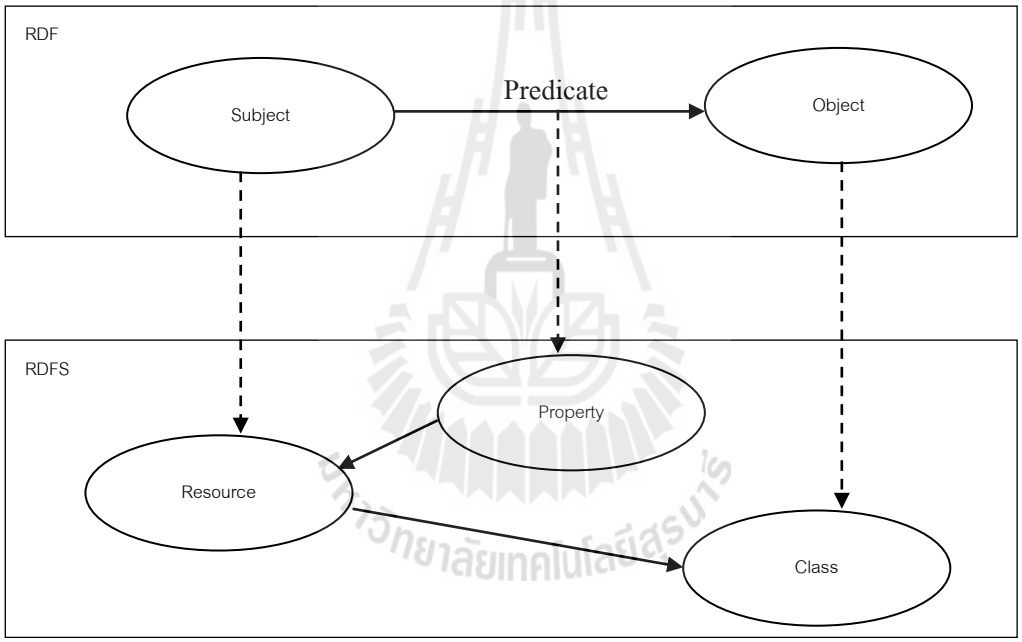
ภาพที่ 2.5 แบบจำลองโครงสร้างข้อมูลของภาษา RDF

จากภาพที่ 2.5 เป็นตัวอย่างภาพที่แสดงส่วนประกอบของภาษาอาร์ดีเอฟ โดยส่วนแรก คือ www.sut.ac.th/db-su เป็นส่วนวัตถุ Site-Owner เป็นส่วนที่แสดงคุณสมบัติของวัตถุ และ Suphakit.N เป็นส่วนที่เป็นค่าของวัตถุ ซึ่งสามารถแปลได้ว่า www.sut.ac.th/db-su เป็นของ Suphakit.N

- **Resource Description Framework Schema (RDFS)**

อาร์ดีเอฟเอส (RDF Schema: RDFS) (Nejdl et al., 2000) เป็นภาษามาตรฐานที่เพิ่มความสามารถในการอธิบายข้อมูลให้มากขึ้น เช่น การอธิบายคลาส การอธิบายการสืบทอดคลาส เนื่องจากโครงสร้างของภาษาอาร์ดีเอฟ นั้นไม่สามารถนิยามการสร้างออนโทโลยีได้ทั้งหมด เป็นเพียงพื้นฐานภาษาสำหรับนิยามออนโทโลยีเบื้องต้นเท่านั้น

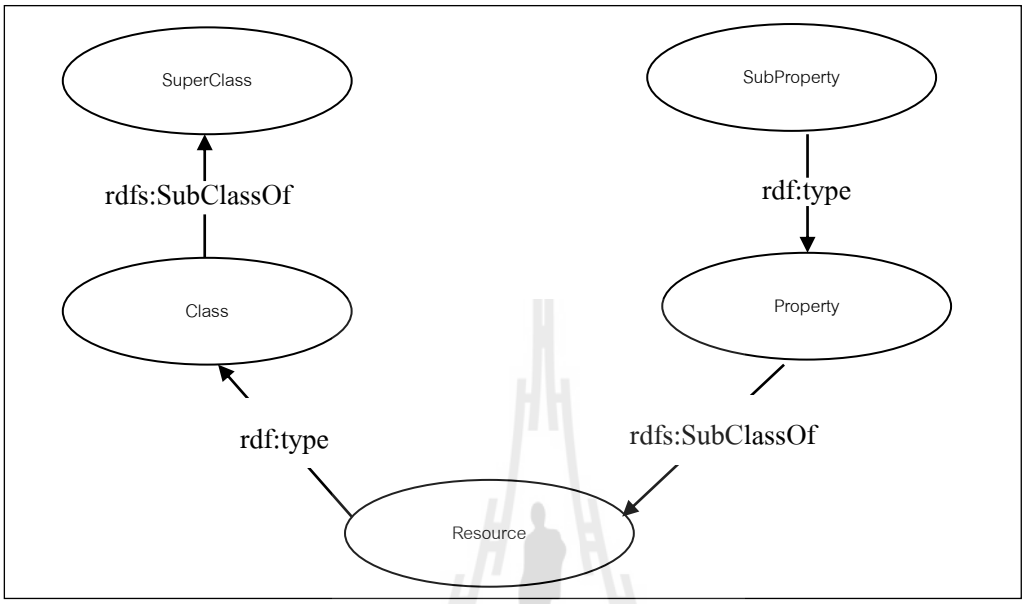
อาร์ดีเอฟเอสเป็นการอธิบายโครงสร้างของเมทาดาทา ซึ่งใช้อธิบาย Resource ได้กำหนด Vocabulary (Guarino, 1998) เพื่อใช้อธิบายโครงสร้างของเมทาดาทา ซึ่งประกอบด้วย คุณสมบัติ และค่าของคุณสมบัติ โดยอธิบายขยายให้อยู่ในรูปของโหนดที่เป็น Property และ Class โดยสามารถสืบทอดเป็น SubProperty และ SubClass ได้ตามลำดับ ซึ่งการอธิบายโครงสร้างของเมทาดาทาสามารถอธิบายได้ ดังภาพที่ 2.6 (Nejdl et al., 2000)



ภาพที่ 2.6 การอธิบายโครงสร้างของเมทาดาทา

จากภาพที่ 2.6 เป็นการนำโครงสร้างเมทาดาทามาใช้อธิบายข้อมูล ซึ่งประกอบไปด้วยคุณสมบัติ และค่าของคุณสมบัติ มากำหนดให้เป็นโหนดเพื่ออธิบาย Resource โดยในส่วนของ Predicate ซึ่งเป็นคุณสมบัติจะถูกกำหนดมาเป็นโหนด Property และส่วนของ Object ซึ่งเป็นค่าของคุณสมบัติจะถูกกำหนดให้เป็นโหนด Class โดยทั้งโหนด Property และโหนด Class จะสามารถสืบทอดเป็น SubProperty และ SubClass ได้ตามลำดับ

เมื่อทำการอธิบายส่วนของเมทาดาทาให้กลายเป็นโหนด Property และโหนด Class รวมถึงมีการสืบทอดเป็น SubProperty และ SubClass จากนั้นจะมีการจัดวางความสัมพันธ์ระหว่างโหนดในรูปแบบของอาร์ดีเอฟเอส ซึ่งมีความสัมพันธ์กันดังภาพที่ 2.7 (Nejdl et al., 2000)



ภาพที่ 2.7 การอธิบายความสัมพันธ์ของ Resource ในรูปแบบของอาร์ดีเอฟเอส

จากภาพที่ 2.7 แสดงให้เห็นถึงการจัดวางความสัมพันธ์ของแต่ละโหนดในรูปแบบของอาร์ดีเอฟเอส โดยโหนด Resource กับโหนด Class มีความสัมพันธ์กันเป็น rdf:type, โหนด Resource กับโหนด Property มีความสัมพันธ์กันเป็น rdfs:SubClassOf, โหนด Class กับ Sub Class มีความสัมพันธ์กันเป็น rdfs:SubClassOf และโหนด Property กับ SubProperty มีความสัมพันธ์กันเป็น rdf:type

ในการสร้างอาร์ดีเอฟเอสนั้น จะกำหนด Vocabulary เพิ่มเข้าไปเพื่ออธิบายโครงสร้างของเมทาดาทาที่อธิบายในโครงสร้างอาร์ดีเอฟเพื่อให้กลายเป็นข้อมูลที่มีความหมายมากขึ้น ดังนั้น อาร์ดีเอฟเอสเป็นภาษาที่เพิ่มความสามารถในการบรรยายข้อมูลให้มากขึ้น เช่น การบรรยายคลาส การบรรยายการสืบทอดคลาส อาร์ดีเอฟเอสเป็นการกำหนดโครงสร้างของเมทาดาทาซึ่งใช้อธิบายความสัมพันธ์กันของแหล่งทรัพยากร ข้อจำกัดของอาร์ดีเอฟเอสประกอบไปด้วยคลาส และคุณสมบัติของคลาส การแสดงคุณสมบัติของคลาสถูกกำหนดโดย rdf:Property อาร์ดีเอฟเอสกำหนดให้มี 4 คุณสมบัติหลักดังนี้ rdfs:subClassOf, rdf:type, rdfs:range, rdfs:domain แต่ด้วยคุณสมบัติเหล่านี้ไม่เพียงพอต่อการบรรยายข้อมูล ให้มีความหมายเชิงตรรกะได้ ดังนั้นจึงได้มีการกำหนดภาษาใหม่ที่ใช้บรรยายโครงสร้าง

ความสัมพันธ์ และความหมายเชิงตรรกะได้โดยภาษาดังกล่าวถูกเรียกว่า “ภาษาโอดับเบิ้ลยูแอล (Web Ontology Lanaguage: OWL)”

- **Web Ontology Language (OWL)**

ภาษาโอดับเบิ้ลยูแอล (Web Ontology Language: OWL) ถูกสร้างโดย W3C Web Ontology Working Group (WebOnt) ภาษาโอดับเบิ้ลยูแอลถูกพัฒนาขึ้นเพื่อเป็นส่วนขยายต่อจากภาษาอาร์ดีเอฟ และสืบทอดมาจากภาษา DAML+OIL ภาษาโอดับเบิ้ลยูแอลจัดเป็นองค์ประกอบหนึ่งในงาน Semantic Web ที่ใช้ในการบรรยายข้อมูลเชิงความหมาย สามารถกำหนดโครงสร้างข้อมูลในลักษณะลำดับชั้น และอธิบายข้อมูล (metadata) ที่มีความสัมพันธ์ในระบบฐานข้อมูลได้ รวมทั้งสามารถรองรับการบรรยายข้อมูลเชิงตรรกะ ชนิดข้อมูล และ ตัวบ่งปริมาณได้ ทำให้ข้อมูลที่ถูกแทนที่นั้นมีความหมายมากยิ่งขึ้น ลักษณะการบรรยายจะอยู่ในรูปคลาส คุณสมบัติของคลาส และความสัมพันธ์ระหว่างคลาส เพื่ออธิบายเอนทิตี้และความสัมพันธ์ต่าง ๆ ที่เกิดขึ้น (Smith et al., 2004)

- *การแบ่งประเภทของภาษาโอดับเบิ้ลยูแอล*

ภาษาโอดับเบิ้ลยูแอลแบ่งออกเป็น 3 ประเภท โดยถูกออกแบบให้เหมาะสำหรับการใช้งานในแต่ละประเภทตามกลุ่มการใช้งานดังนี้

1) OWL Life

ออกแบบมาเพื่อสนับสนุนการใช้งานเบื้องต้น จะมีการกำหนดโครงสร้างในรูปแบบลำดับชั้นและมีการบังคับใช้คุณสมบัติพื้นฐานในการกำหนดโครงสร้างข้อมูล ถูกออกแบบมาให้ง่ายในการพัฒนาและมีการเตรียมฟังก์ชันการใช้งานต่าง ๆ เพื่อสำหรับเริ่มใช้งานการเขียนภาษาโอดับเบิ้ลยูแอลได้

2) OWL DL

ออกแบบมาเพื่อสนับสนุนการอธิบาย Logic Business Segment โดยใน OWL DL จัดให้มีคุณสมบัติที่เหมาะกับการใช้งานด้านฐานข้อมูล และการแทนความรู้ที่ตั้งอยู่บนพื้นฐานของการอธิบายด้วยเหตุผลทางตรรกะ OWL DL สามารถบรรยายข้อมูลและโครงสร้างข้อมูล ในรูปแบบโครงสร้างภาษาโอดับเบิ้ลยูแอลด้วยข้อจำกัดของคลาสและคุณสมบัติของคลาสได้

3) OWL FULL

ออกแบบมาเพื่อสนับสนุนผู้ใช้งานที่ต้องการความครบถ้วนและมีโครงสร้างภาษาที่สมบูรณ์แบบ โดย OWL Full จะมีการผสมผสานกันระหว่างภาษาโอดับเบิ้ลยูแอลและอาร์ดีเอฟเอส ผู้ใช้งานสามารถบรรยายข้อมูลในรูปแบบอาร์ดีเอฟเอสได้อย่างอิสระทั้ง OWL DL และ OWL Full ต่างก็สนับสนุนเซตของภาษาโอดับเบิ้ลยูแอลด้วยกันทั้งนั้น แต่มีข้อจำกัดของคุณลักษณะบางอย่างที่แตกต่างกันบนพื้นฐานของอาร์ดีเอฟเอส โดย OWL Full จะมีการผสมผสานกันระหว่างภาษาโอดับเบิ้ลยูแอล

และอาร์ดีเอฟเอสโดยไม่มีการบังคับในส่วนการแบ่งคลาส การกำหนดคุณสมบัติ และค่าของข้อมูล ส่วน OWL DL จะมีข้อบังคับในการใช้อาร์ดีเอฟในการกำหนดคลาส การกำหนดคุณสมบัติ และค่าของข้อมูล เป็นต้น

- *โครงสร้างของภาษาโอดับเบิ้ลยูแอล*

ภาษาโอดับเบิ้ลยูแอลมีการบรรยายข้อมูลแบบผสมผสานกันระหว่างการใช้ไวยกรณ์ของอาร์ดีเอฟ, อาร์ดีเอฟเอส และเอ็กซ์เอ็มแอลซึ่งแบ่งตามประเภทของการใช้งาน ผลลัพธ์ที่ได้จะอยู่ภายใต้รูปแบบของ RDF Triples เอกสารของภาษาโอดับเบิ้ลยูแอล จะประกอบด้วยกลุ่มข้อมูลของ Namespace, Ontology Headers, Class, Property และรายละเอียดของข้อมูลต่าง ๆ ไฟล์นามสกุลที่ใช้ในการสร้างเอกสารเป็นไฟล์นามสกุล .rdf หรือ .owl สามารถแยกรายละเอียดส่วนประกอบต่างๆ ได้ดังนี้

1) Namespace

การกำหนด Namespace จะประกาศไว้ที่ส่วนเริ่มต้นของเอกสาร เพื่อเป็นการกำหนดกลุ่มในการอ้างอิงข้อมูล เอกสารโอดับเบิ้ลยูแอลที่ถูกสร้างจะขึ้นอยู่กับโครงสร้างที่ถูกนิยามด้วยอาร์ดีเอฟ, อาร์ดีเอฟเอส และชนิดข้อมูลของอาร์ดีเอฟเอส การเขียน Namespace จะประกาศไว้ภายใต้คำสั่งของ rdf:RDF syntax แสดงตัวอย่างได้ดังตัวอย่าง

ตัวอย่าง รหัสต้นฉบับการกำหนด Namespace

```
<rdf:RDF 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#"
xmlns:xsd="http://www.w3.org/2000/10/XMLSchema#"
```

2) Ontology Headers

แสดงการอธิบายรายละเอียดเบื้องต้นของออนโทโลยี ว่าเป็นโครงสร้างข้อมูลเกี่ยวกับอะไร ภายใต้อิลิเมนต์ <owl:Ontology rdf:about=""> ประกอบด้วยคำสั่ง <owl:versionInfo> ใช้แสดงรุ่นของข้อมูลที่ทำการสร้าง คำสั่ง <rdfs:comment> ใช้แสดงส่วนการอธิบายหมายเหตุของข้อมูล คำสั่ง <owl:imports rdf:resource=" "> ใช้แสดงการอ้างอิงเกี่ยวกับแหล่งข้อมูลว่ามาจากที่ใด ตัวอย่างการแทนค่าใช้งานแสดงได้ดังตัวอย่าง

ตัวอย่าง รหัสต้นฉบับโครงสร้าง Ontology Headers

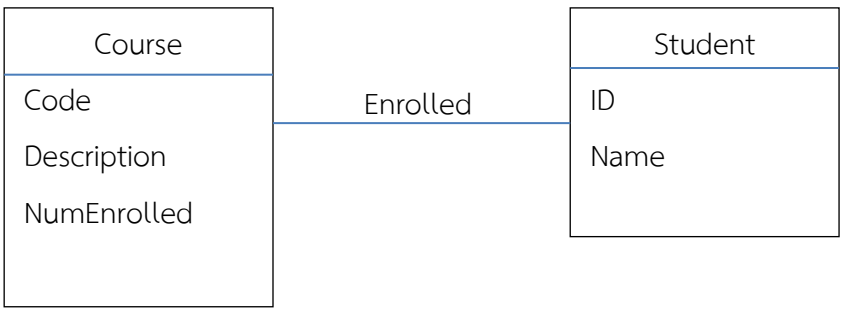
```
<owl:Ontology rdf:about="">
<owl:versionInfo>V 1.17 2003/02/26 </owl:versionInfo>
<rdfs:comment>An example ontology</rdfs:comment>
<owl:imports rdf:resource="http://www.example.org/owl"/></owl:Ontology>
```

- Basic Element

อิลิเมนต์พื้นฐานของ OWL Ontology ส่วนมากจะเกี่ยวข้องกับคลาส, คุณสมบัติของคลาส, ตัวแทนข้อมูล (instances ของคลาส) และความสัมพันธ์ระหว่างตัวแทนข้อมูลต่างๆ มีรายละเอียดดังต่อไปนี้

1) คลาสและคลาสย่อย

คลาสเป็นเค้าโครงที่ใช้กำหนดกลุ่มของข้อมูล หรือ Object ซึ่งเป็นแหล่งข้อมูลที่มีคุณลักษณะที่คล้ายคลึงกัน โดยทุก ๆ OWL คลาสจะมีความสัมพันธ์กับเซตของข้อมูลที่มีลักษณะเฉพาะหรือเรียกว่า Class Extension ซึ่งเป็นลักษณะของการสืบทอดคลาส โดยลักษณะเฉพาะ (Individual) ภายใน Class Extension ถูกเรียกว่าตัวแทนข้อมูล โดยหลักการพื้นฐานของการบรรยายคลาสจะมองเป็นโครงสร้างลำดับชั้นมีการสืบทอดคุณสมบัติของคลาส ในโครงสร้างของภาษาโอดับเบิ้ลยูแอลกล่าวได้ว่าทุกๆ สิ่งที่มีลักษณะเฉพาะอย่างทั้งหลายจะเป็นสมาชิกของคลาส owl:Thing โดยกำหนดให้ owl:Thing เป็นคลาสใหญ่ที่สามารถครอบคลุม ทุกคลาสข้อมูลได้ ดังนั้นไม่ว่าผู้ใช้งานกลุ่มใดสร้างคลาสขึ้นมา จะเสมือนว่าเป็นสมาชิกอยู่ภายใต้คลาส owl:Thing นี้ และในทางตรงข้ามคลาส owl:Nothing ใช้ในการกำหนดคลาสที่ไม่มีสมาชิกภายในคลาสนั้น หรือเป็นคลาสว่าง ตัวอย่างการสร้างคลาส จากฐานข้อมูลเชิงสัมพันธ์ของระบบลงทะเบียน โดยอ้างอิงเพียงบางตารางข้อมูลที่เกี่ยวข้อง เพื่อใช้เป็นตัวอย่างการแทนข้อมูลได้ ดังแสดงในภาพที่ 2.8 (Hart et al., 2004)



ภาพที่ 2.8 ตัวอย่างตารางข้อมูลของระบบลงทะเบียน

จากภาพที่ 2.8 แสดงตัวอย่างตารางข้อมูลของระบบลงทะเบียน ประกอบด้วยตารางข้อมูลที่มีความสัมพันธ์กัน 2 ตาราง คือ ตารางหลักสูตร (Course) และตารางข้อมูลนักศึกษา (Student) ซึ่ง 2 ตารางนี้มีความสัมพันธ์ที่มีชื่อว่าลงทะเบียน (Enrolled) ในตารางข้อมูลจะประกอบด้วยคุณสมบัติของตาราง ซึ่งแสดงถึงองค์ประกอบของตารางว่าประกอบด้วยข้อมูลอะไรบ้าง เช่น ตารางหลักสูตร คุณสมบัติ ของตารางประกอบด้วย รหัสหลักสูตร (Code) คำอธิบายหลักสูตร (Description) และจำนวนผู้ลงทะเบียน (NumEnrolled) ส่วนตารางข้อมูลนักศึกษาคุณสมบัติของตารางประกอบด้วย รหัสนักศึกษา (Id) และชื่อนักศึกษา (Name) เป็นต้น ดังนั้นรูปแบบของการเขียนคลาส สามารถกำหนดการประกาศใช้งานคลาสแบบเฉพาะเจาะจง สามารถประกาศชื่อคลาสได้ตามที่ต้องการดังตัวอย่าง

ตัวอย่าง การประกาศชื่อคลาส

```
<owl:Class rdf:ID="Course"/>
<owl:Class rdf:ID="Student"/>
```

จากตัวอย่าง การกำหนดชื่อคลาสอยู่ภายใต้ syntax rdf:ID="Course" ภายใต้เครื่องหมาย " " จะแทนชื่อคลาสที่ถูกกำหนดขึ้น การแทนชื่อคลาสในภาษา OWL สามารถเขียนได้หลายรูปแบบ เช่น #Course เขียนแทนด้วย syntax rdf:resource="#Course" โครงสร้างที่เป็นพื้นฐานของคลาส คือ rdfs:subClassOf โดยจะมีความสัมพันธ์กับการเฉพาะเจาะจงกับคลาสโดยทั่วไปดังนี้ ถ้า X เป็น subclass ของ Y ดังนั้น ทุก ๆ instance ของ X จะเป็น instance ของ Y ด้วย ซึ่ง rdfs:subClassOf เป็นความสัมพันธ์แบบ transitive ถ้า X เป็น subclass ของ Y และ Y เป็น subclass ของ Z ดังนั้น X เป็น subclass ของ Z ด้วย ดังตัวอย่างการแทนค่าข้อมูล โดยกำหนดให้ Subject เป็น subclass ของ Course แสดงได้ดังตัวอย่าง

ตัวอย่าง การประกาศชื่อคลาสย่อย

```
<owl:Class rdf:ID="Subject">
<rdfs:subClassOf rdf:resource="#Course"/>
</owl:Class>
```

การนิยามคลาสจะประกอบด้วย 2 ส่วนได้แก่ ชื่อคลาส ที่ใช้แทนการอ้างอิงการใช้งานคลาสนั้น และการแสดงรายละเอียดหรือข้อจำกัดต่าง ๆ ส่วนของการอธิบายสมาชิกภายในคลาสนั้นทำได้โดยการประกาศถึงคุณลักษณะเฉพาะของสมาชิกภายในคลาส แสดงได้ดังตารางที่ 2.2

ตารางที่ 2.2 ข้อมูลสมาชิกภายในคลาส

Classifier	code	Description	NumEnrolled
Course	INFS3101	Ontology and Semantic Web	0

จากตารางที่ 2.2 แสดงข้อมูลสมาชิกภายในคลาส Course ข้อมูลที่ใช้แสดงถึงข้อมูลเฉพาะ คือ code เขียนบรรยายในรูปแบบภาษา OWL ได้ดังนี้ <owl:Thing rdf:ID=" INFS3101"/> (Hart et al., 2004)

2) คุณสมบัติของคลาส

ภาษา OWL ได้แบ่งคุณสมบัติออกเป็น 2 ประเภทได้แก่ Object Properties และ Datatype Properties ดังแสดงในตารางที่ 9 การกำหนดคุณสมบัติในภาษา OWL สนับสนุนตามโครงสร้างของคุณสมบัติต่างรูปแบบ ดังแสดงในตารางที่ 2.3 และตารางที่ 2.4

ตารางที่ 2.3 ประเภทของคุณสมบัติ

ประเภทของคุณสมบัติ	นิยาม
1. Object Properties	แสดงความสัมพันธ์ระหว่าง instance ของคลาสที่เกี่ยวข้องกัน 2 คลาส
2. Datatype Properties	แสดงความสัมพันธ์ระหว่าง instance ของคลาสและ RDF literals รวมถึงการ แสดงชนิดข้อมูล XML Schema datatypes

ตารางที่ 2.4 โครงสร้างของคุณสมบัติในรูปแบบต่าง ๆ

RDF Schema:	rdfs:subPropertyOf, rdfs:domain และ rdfs:range
Relations to other properties:	owl:equivalentProperty และ owl:inverseOf
global cardinality constraints:	owl:FunctionalProperty และ owl:InverseFunctionalProperty
logical property characteristics:	owl:SymmetricProperty และ owl:TransitiveProperty

3) คุณสมบัติและชนิดข้อมูล

การเขียนโครงสร้างของภาษาโอดับเบิ้ลยูแอลประกอบด้วยชนิดของข้อมูลที่ต้องการใช้งาน ซึ่งการประกาศใช้งานจะผ่านคำสั่งในรูปแบบ XML Schema datatypes โดยการประกาศใช้จะอ้างอิงตามชนิดข้อมูลที่ใช้งานจริง และมีการประกาศการใช้งาน URI ที่มีการอ้างอิงการใช้งานชนิดข้อมูลนั้นๆ ได้แก่ http://www.w3.org/2001/XMLSchema#NAME โดยที่ NAME แทนการอ้างอิงประเภทของชนิดข้อมูล เช่น URI="http://www.w3.org/2001/XMLSchema#string"/> แสดงถึงการใช้งานชนิดข้อมูลประเภท string ชนิดข้อมูลที่อ้างอิงตาม XML Schema datatypes แสดงได้ดังตารางที่ 2.5

ตารางที่ 2.5 ชนิดข้อมูลที่อ้างอิงตาม XML Schema datatypes

กลุ่มชนิดข้อมูล	**คำสั่งที่ใช้งาน**
ประเภท string	xsd:string, xsd:normalizedString, xsd:token, xsd:language, xsd:NMTOKEN, xsd:Name, xsd:NCName, xsd:hexBinary, xsd:base64Binary, xsd:anyURI.
ประเภท boolean	xsd:boolean
ประเภท numerical	xsd:decimal, xsd:float, xsd:double, xsd:integer, xsd:positiveInteger. xsd:nonPositiveInteger, xsd:negativeInteger, xsd:nonNegativeInteger, xsd:long, xsd:int, xsd:short, xsd:byte, xsd:unsignedLong, xsd:unsignedInt, xsd:unsignedShort, xsd:unsignedByte
ประเภท time	xsd:dateTime, xsd:time, xsd:date, xsd:gYearMonth, xsd:gYear, xsd:gMonthDay, xsd:gDay, xsd:gMonth.

4) ข้อกำหนดของคุณสมบัติ

เพื่อจัดการความหลากหลายของข้อมูล ภาษาโอดับเบิ้ลยูแอลได้ตั้งข้อกำหนดการใช้งานคุณสมบัติ และลักษณะพิเศษของการบรรยายคลาส โดยอธิบายคลาสในรูปแบบ anonymous class รวมถึงคลาสทุกคลาสที่เป็น individual class ได้ถูกกำหนดให้เป็นข้อจำกัดในการใช้งานคุณสมบัติ

โครงสร้างภาษาโอดับเบิ้ลยูแอลแบ่งข้อกำหนดของคุณสมบัติเหล่านี้ได้ 2 ประเภทได้แก่ value constraints และ cardinality constraints ดังแสดงในตารางที่ 2.6

ตารางที่ 2.6 ประเภทข้อกำหนดของคุณสมบัติ

ประเภท	คำสั่งที่ใช้งาน
1. value constraints	owl:allValuesFrom, owl:someValuesFrom, owl:hasValue
2. cardinality constraints	owl:maxCardinality, owl:minCardinality, owl:cardinality

ตัวอย่างการแทนค่าข้อกำหนดของคุณสมบัติโดยคำสั่ง owl:allValuesFrom เป็นการจำกัดความต้องการสำหรับทุก ๆ instance ของคลาสให้มีคุณสมบัติที่เป็นลักษณะเฉพาะเท่านั้นซึ่งค่าของคุณสมบัตินี้ทุก ๆ สมาชิกภายในคลาสจะถูกกำหนดภายใต้คำสั่ง owl:allValuesFrom เช่น การนิยามว่ามนุษย์ทุกคนจะต้องมีผู้ให้กำเนิดดังตัวอย่าง

ตัวอย่าง การกำหนดคุณสมบัติคลาสด้วยคำสั่ง owl:allValuesFrom

```
<owl:Restriction>
<owl:onProperty rdf:resource="#hasParent" />
<owl:allValuesFrom rdf:resource="#Human" />
</owl:Restriction>
```

ตัวอย่างคำสั่ง owl:someValuesFrom แสดงความหมายว่าโดยอย่างน้อย 1 ค่าของคุณสมบัติจะต้องเกี่ยวข้องกับ instance ของการบรรยายคลาส หรือค่าข้อมูลภายในขอบเขตของข้อมูล เช่น การนิยามว่าโดยทั่วไปผู้ให้กำเนิดอาจมีชีวิตอยู่ทั้งสองคน หรืออาจมีชีวิตอยู่เพียงบิดาหรือมารดาดังตัวอย่าง

ตัวอย่าง การกำหนดคุณสมบัติคลาสด้วยคำสั่ง owl:someValuesFrom

```
<owl:Restriction>
<owl:onProperty rdf:resource="#hasParent" />
<owl:someValuesFrom rdf:resource="#Physician" />
</owl:Restriction>
```

สรุปได้ว่าภาษาโอดับเบิ้ลยูแอลได้รับการพัฒนาขึ้นเพื่อปิดจุดด้อยของอาร์ดีเอฟโดยการพัฒนาการออกแบบอาร์ดีเอฟได้ถูกเพิ่มเติมด้วยภาษาโอดับเบิ้ลยูแอล ซึ่งเป็นเครื่องมือสำหรับการอธิบายความสัมพันธ์ระหว่างคลาส, การกำหนดคุณลักษณะ, จำนวนสมาชิกและการตั้งค่าที่เกี่ยวข้องกับคุณสมบัติ ฯลฯ ในทางปฏิบัติ การพัฒนาเว็บเชิงความหมายไม่สามารถเขียนได้โดยใช้ภาษาอาร์ดีเอฟหรือภาษาโอดับเบิ้ลยูแอลได้ด้วยมือ โดยมีซอฟแวร์หลายๆซอฟแวร์ที่ใช้สำหรับการพัฒนาการเขียนภาษาอาร์ดีเอฟและภาษาโอดับเบิ้ลยูแอลได้อย่างอัตโนมัติ

2.6 การจัดการฐานความรู้ด้วยตัวแบบออนโทโลยี

ออนโทโลยี (Ontology) เป็นศาสตร์ที่ใช้ในการพัฒนาความรู้ที่มีลักษณะเชิงเนื้อหา (Content base) เป็นความรู้ในเชิงประสบการณ์ ปัญหาและการแก้ไข ในปัจจุบันออนโทโลยีได้ถูกนำไปใช้ในงานวิจัยหลายๆ ด้าน เช่น ด้านปัญญาประดิษฐ์ (AI), วิศวกรรมความรู้ (Knowledge Engineering), ประยุกต์ใช้ในการจัดการภาษาธรรมชาติ (Natural Language Processing: NLP) โดยมีผู้ให้คำนิยามของออนโทโลยีไว้หลายแบบ ดังต่อไปนี้

ออนโทโลยี (วิลาศ วูวงศ์ และคณะ, 2548; Fensel et al., 2001) คือ การนิยามหรือกำหนดรูปแบบโครงสร้างของสิ่งที่เราสนใจให้มีความหมายตามขอบเขตขององค์ความรู้ หรือออนโทโลยี คือ นิยามที่เป็นทางการและประกาศอย่างชัดแจ้งของ คำศัพท์ แนวคิด หรือแบบจำลองที่ใช้ร่วมกัน ออนโทโลยีจะบรรยายในลักษณะของคลาส (Class) ความสัมพันธ์ระหว่างคลาส และคุณสมบัติของคลาส (Properties) โดยรูปแบบของการบรรยายจะขึ้นอยู่กับภาษาที่ใช้แทนข้อมูลเชิงความหมาย เช่น เอ็กซ์เอ็มแอล, อาร์ดีเอฟ และภาษาโอดับเบิ้ลยูแอล เป็นต้น โดยแต่ละภาษามีความสามารถและข้อจำกัดในการบรรยายข้อมูลได้แตกต่างกันไป เช่น ภาษาเอ็กซ์เอ็มแอลเหมาะสำหรับใช้อธิบายข้อมูลในส่วนเนื้อหาทำให้ภาษาดังกล่าวไม่สามารถนำมาอธิบายความหมายของข้อมูลได้ ภาษาอาร์ดีเอฟและอาร์ดีเอฟเอสเป็นภาษาที่ใช้อธิบายข้อมูล ความสัมพันธ์ของข้อมูล และโครงสร้าง ภาษานี้ได้เพิ่มความสามารถที่ใช้บรรยายข้อมูลแต่ยังมีข้อจำกัดด้านการบรรยายข้อมูลเชิงตรรกะ ภาษาโอดับเบิ้ลยูแอลเป็นภาษาที่ใช้บรรยายข้อมูลเชิงความหมายและตรรกะสามารถใช้บรรยายโครงสร้างความสัมพันธ์แบบลำดับชั้นในฐานข้อมูลเชิงความสัมพันธ์ได้ ดังนั้น ภาษาโอดับเบิ้ลยูแอลจึงเป็นภาษาที่มีความสามารถในการบรรยายข้อมูลเชิงความหมาย และบรรยายโครงสร้างความสัมพันธ์ของระบบได้ดีกว่าภาษาอื่น ๆ

ออนโทโลยีคือทฤษฎีทางตรรกะที่ใช้กำหนดความหมายของการอธิบายคำศัพท์ (Vocabulary) ซึ่งมีการกำหนดการอธิบายความหมายนั้นด้วยรูปแบบอย่างเป็นทางการ (Guarino, 1998) ออนโทโลยีถูกนำมาใช้ในด้านวิทยาการคอมพิวเตอร์เพื่อให้มีการกำหนดคุณลักษณะของสิ่งที่จะมาแบ่งปันหรือใช้ร่วมกันให้เป็นรูปแบบและมีความชัดเจน (Gruber, 1993) การวิจัยเกี่ยวกับออนโทโลยีในปัจจุบันนี้มุ่งเน้นของการแลกเปลี่ยนข้อมูลระหว่างกันและการนำมาใช้ใหม่ ออนโทโลยีเป็นการทำโครงสร้างและ

การนำเสนอความรู้ในระบบการจัดระบบความรู้ ซึ่งเป็นเสมือนเครื่องมือที่มีคุณลักษณะเป็นภาษาดรรชนีที่ช่วยในการจัดการและเข้าถึงเนื้อหา (Subject Access) โดยเป็นภาษาควบคุม (Controlled Vocabularies) ที่ประกอบด้วยคำศัพท์ (Vocabularies /Terms) และไวยากรณ์ (Syntax) หรือ กฎ (Rules) เพื่อประโยชน์ในการกำหนดรูปแบบคำศัพท์ที่เหมาะสม (Broughton, 2006)

ออนโทโลยีมีการกำหนดโครงสร้างความรู้ที่ชัดเจนของสิ่งที่เกี่ยวข้องกับแนวคิด (Concepts) เพื่อนำมาใช้ในการบรรยายหรือแสดงความคิดในขอบเขตความรู้เรื่องใดเรื่องหนึ่ง (Domain) สามารถแสดงให้เห็นถึงคุณสมบัติที่เกี่ยวข้องกับสิ่งนั้น ๆ ซึ่งรูปแบบโดยทั่วไปของออนโทโลยีจะอยู่ในรูปโครงสร้างข้อมูลแบบมีลำดับชั้น (Hierarchical Data Structure) และมีการจัดกลุ่มแนวคิดแบบแทกโซโนมี (Taxonomy) ในลักษณะ parent-child ดังนั้นโครงสร้างออนโลยีจึงประกอบด้วย การกำหนดนิยามความหมายหรือการอธิบายเรื่องราวโดยใช้แนวคิด (Concept) คุณสมบัติ (Properties) ความสัมพันธ์ (Relationships) ตรรกะการแปลงความสัมพันธ์ (Axioms) และคำศัพท์หรือตัวอย่างข้อมูล (Instances) (Duineveld et al, 2000)

1. แนวคิด (Concept) หมายถึงแนวคิดในขอบเขตความรู้เรื่องใดเรื่องหนึ่ง ซึ่งสามารถเป็นอะไรก็ได้ที่เกี่ยวกับสิ่งที่ถูกกล่าวถึงและสามารถอธิบายรายละเอียดได้

2. คุณสมบัติ (Properties, Slots, Roles, Functions) คุณสมบัติต่าง ๆ ที่เกี่ยวข้องสัมพันธ์กับแนวคิด (Concept) เพื่อนำมาใช้ในการอธิบายแนวคิด (Jacob, 2003) เช่น ราคาของรถมือสอง (Price-of a-Used-Car) และระยะทางการขับ (Mileage) เป็นต้น

3. ความสัมพันธ์ (Relations) เป็นการนำเสนอประเภทของการปฏิสัมพันธ์ (Interaction) ระหว่างแนวคิดกับแนวคิดในขอบเขตความรู้ที่สนใจ และมีการกำหนดแบบแผนความสัมพันธ์ มีลักษณะต่าง ๆ ได้แก่ ความสัมพันธ์ตามลำดับชั้นแบบแทกโซโนมี (Taxonomy) (Subclass-of หรือ Is_a hierarchy) และความสัมพันธ์แบบ Mereology (Part-of) เช่น Trip Is_a Travel Product

4. แอกเซียม (Axiom) เป็นตรรกะของการแปลงความสัมพันธ์ระหว่างแนวคิด (Concepts) กับคุณสมบัติ (Properties) เพื่อการแปลงความหมายที่ถูกต้อง

5. คำศัพท์หรือตัวข้อมูล (Instances) เป็นคำศัพท์ที่มีการกำหนดนิยามความหมายร่วมกันทั้งหมดในออนโทโลยี

- **ประโยชน์ของออนโทโลยี**

ออนโทโลยีสามารถนำมาประยุกต์ใช้ในการทำงานของระบบงานต่างๆ โดยพิจารณาได้จากวัตถุประสงค์ของการนำไปใช้ในงาน ดังต่อไปนี้

1) ระบบสืบค้นข้อมูล ออนโทโลยีมีส่วนช่วยในการขยายคำค้น ทำให้ค้นหาและเข้าถึงข้อมูลที่ตรงตามที่ผู้ใช้งานต้องการจริง ด้วยการพิจารณาความหมายของสิ่งต่าง ๆ ในโดเมนเช่น ระบบงานการค้นหายา (Drug Discovery) ระบบค้นหาข้อมูลทางพันธุกรรม (GeneOntology) (Berners-Lee, 2006)

2) ออนโทโลยีต่างจากฐานข้อมูล ในด้านของการเก็บรายละเอียดต่าง ๆ เช่น บทบาท (Role) ความสัมพันธ์ (Relation) ในขณะที่ฐานข้อมูลเป็นเพียงแหล่งรวบรวมข้อมูล หากใช้ ออนโทโลยีช่วยในการค้นหาข้อมูล จะช่วยทำให้ได้ข้อมูลข่าวสารที่ถูกต้องยิ่งขึ้น

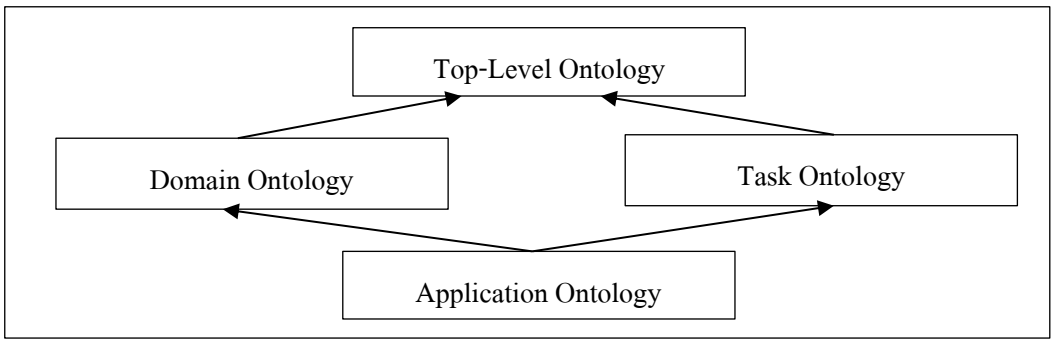
3) ระบบงานที่ต้องการผนวกความรู้ (Knowledge Integration) เพื่อสร้างความเข้าใจเบื้องต้นของความรู้ระหว่างโดเมน

4) ระบบงานที่มีความต้องการนำข้อมูลกลับมาใช้อีก (Reuse) เช่น ระบบงานหนึ่งอาจจะกำหนดออนโทโลยีเพื่อใช้งาน และมีการอนุญาตให้ระบบงานอื่น สามารถนำไปใช้งานได้อีกโดยที่ระบบงานที่นำไปใช้งานภายหลัง อาจกำหนดข้อมูลเชิงความหมายเพิ่มเติมได้โดยไม่ขัดแย้งกับข้อมูลเชิงความหมายของออนโทโลยีที่ถูกนำข้อมูลกลับมาใช้อีก

5) ระบบงานที่มีความต้องการในการอธิบายข้อมูลที่ไม่มีการเปลี่ยนแปลง (Berners-Lee, 2006) เช่นใช้ในการอธิบายความรู้ในโดเมนแบบสถิตย์ (Static Domain Knowledge) ของระบบองค์ความรู้ (Knowledge-Based System)

- **ประเภทของออนโทโลยี**

ออนโทโลยีสามารถออกแบบได้ในแง่มุมต่างๆ ขึ้นกับวัตถุประสงค์ของการใช้งานและความเหมาะสมในการอธิบายข้อมูล โดยพิจารณาประเภทของออนโทโลยี ได้ดังภาพที่ 2.9 (Hunter et al, 2002)



ภาพที่ 2.9 ประเภทออนโทโลยี

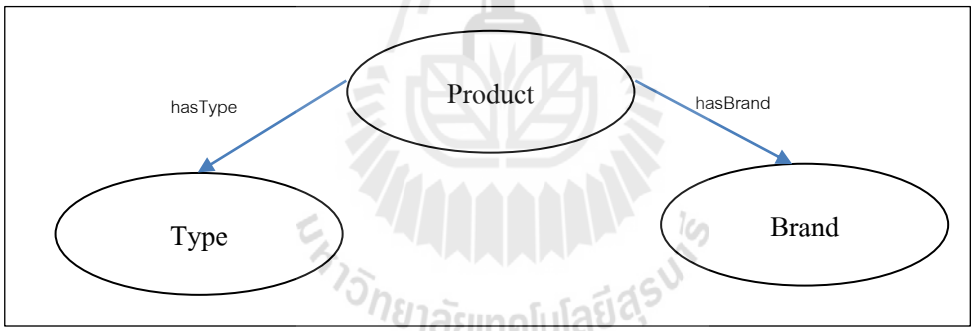
จากภาพที่ 2.9 ประกอบไปด้วย Top-level Ontology (upper ontology), Domain Ontology,Task ontology, Application Ontology

1. ออนโทโลยีระดับบน (Top-level Ontology or Upper Ontology) เป็นออนโทโลยีที่ประกอบด้วยเบสคลาส (Base Class) และกำหนดคุณสมบัติเพื่ออธิบายคลาส หรือกำหนดความสัมพันธ์ระหว่างคลาส โดยสามารถนำไปใช้งานได้ในโดเมนทั่วไป(Generic domain)

2. ออนโทโลยีสำหรับกิจกรรม (Task Ontology) เป็นออนโทโลยีที่พัฒนาขึ้นเพื่อตอบสนองการทำงานของกิจกรรมย่อยๆ โดยอาศัยการถ่ายทอดคุณลักษณะเฉพาะของกิจกรรมจากออนโทโลยีระดับบน

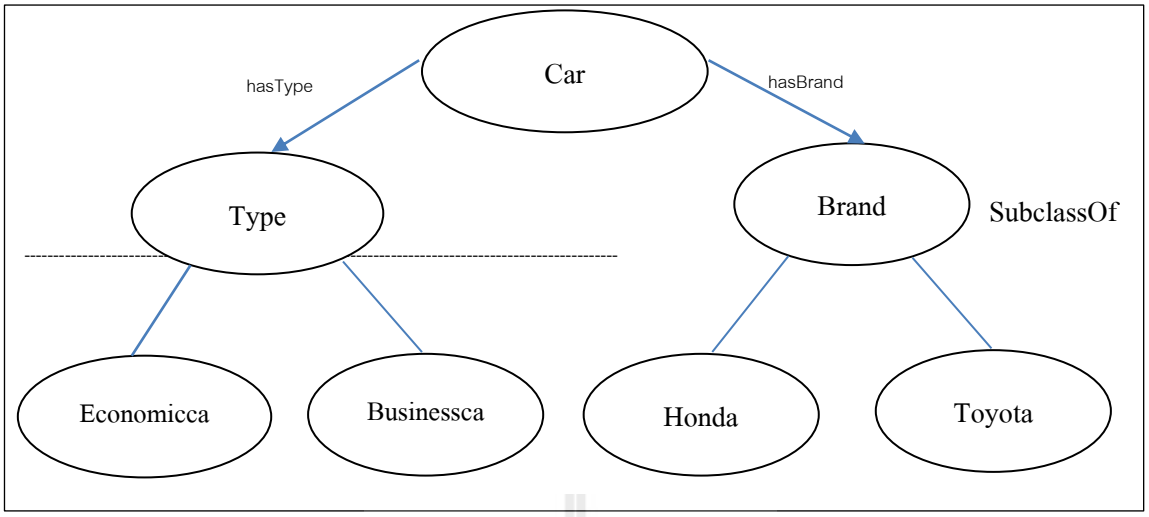
3. ออนโทโลยีสำหรับโดเมน (Domain Ontology) เป็นออนโทโลยีที่ตอบสนองต่อโดเมนโดยอาศัย การถ่ายทอดคุณลักษณะเฉพาะของโดเมนจากออนโทโลยีระดับบน

4. ออนโทโลยีระดับโลคอล (Application Ontology or Local Ontology) เป็นออนโทโลยีที่ถูกจำกัดการใช้งานในโดเมนที่มีความจำเพาะเจาะจง (Specific domain)



ภาพที่ 2.10 Upper Ontologyของสินค้า

จากภาพที่ 2.10 แสดงความหมายของออนโทโลยีของสินค้า ซึ่งสามารถนำไปประยุกต์ใช้งานในโดเมนที่ต่างกันได้ เช่น อาจนำไปใช้ในการอธิบายสินค้าประเภท รถยนต์ โดยภาพที่ 2.11 แสดงตัวอย่างของออนโทโลยีระดับโลคอลเมื่อถูกนำไปใช้อธิบายสินค้าประเภทรถยนต์



ภาพที่ 2.11 Local Ontology ของรถยนต์

จะเห็นว่าออนโทโลยีรถยนต์ได้ถูกออกแบบมาเพื่อแสดงการแบ่งชนิดของรถยนต์ (Type) และแบรนด์ (Brand) ซึ่งเป็นคลาสย่อยระดับที่ 1 และมีคลาสย่อยระดับที่ 2 ประกอบด้วย คลาส Economic car, คลาส Business Car, คลาส Honda และคลาส Toyota เป็นต้น ตารางที่ 2.7 อธิบายความหมายอคลาสระดับต่าง ๆ เหล่านี้

ตารางที่ 2.7 อธิบายคลาสที่พบในออนโทโลยีรถยนต์

ระดับคลาส	คำอธิบาย
คลาสข้อมูล	คลาสหลักและคลาสย่อย
คลาสหลัก	คลาสรถยนต์ (Car)
คลาสย่อยระดับ 1	คลาส Type, คลาส Brand
คลาสย่อยระดับ 2	คลาส Economic car, คลาส Business Car, คลาส Honda, คลาส Toyota

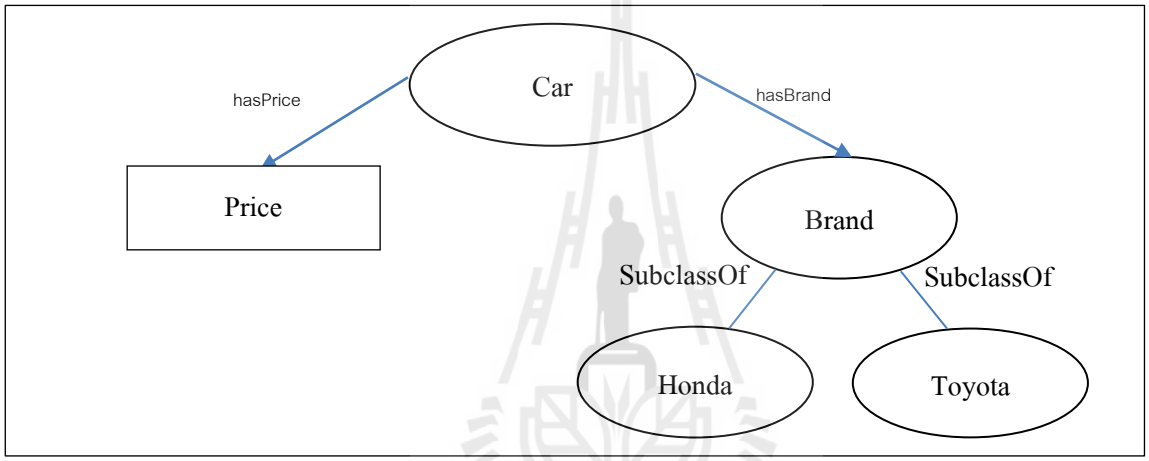
- **ส่วนประกอบหลักของออนโทโลยี**

1) การกำหนดข้อมูลเค้าร่างสำหรับการอธิบายข้อมูลเชิงความหมาย คือการกำหนด ออนโทโลยีระดับบนเพื่อเป็นโมเดลแสดงโครงสร้างการอธิบายข้อมูลเชิงความหมาย ซึ่งจะต้องกำหนดคลาส (Class) พร็อพเพอร์ตี้ (Property) และเงื่อนไข (Restriction) สำหรับการอธิบายเค้าโครงร่างข้อมูลแม่แบบสำ หรับการอธิบายข้อมูลเชิงความหมาย ซึ่งในการอธิบายข้อมูลประกอบด้วย

1. คลาส (Class) หรือ คอนเซ็ปต์ (Concept) หรือ (Category) เป็นตัวแสดงถึงความรู้ที่เราสนใจ และอธิบายได้ว่าคลาสต่าง ๆ บรรจุอะไรไว้ในโดเมน คลาสเป็นส่วนที่จะต้องพิจารณาอย่างละเอียดในการพัฒนาออนโทโลยี

2. พร็อพเพอร์ตี้ (Property) หรือความสัมพันธ์ (Relation) หรือสล็อต (Slot) แสดงถึงการกำหนดความสัมพันธ์ (Relation) หรือคุณลักษณะของคลาส เพื่อเชื่อมโยงระหว่างคลาส ด้วยการระบุค่าพร็อพเพอร์ตี้ ที่สามารถกำหนดได้ด้วยการประกาศให้เป็นค่าคงที่

จากภาพที่ 2.12 เป็นการแสดงตัวอย่างข้อมูลเค้าร่างที่อธิบายข้อมูลรายละเอียดของคลาสรถยนต์ โดยมีการประกาศคลาสและพร็อพเพอร์ตี้ดังแสดงในตารางที่ 2.8

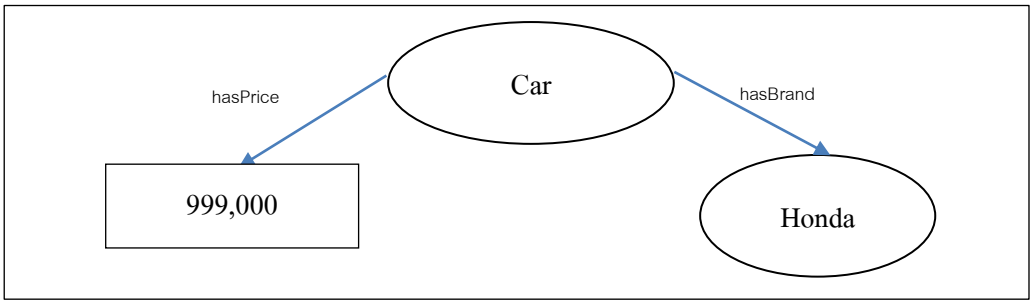


ภาพที่ 2.12 ตัวอย่างข้อมูลเค้าร่างอธิบายข้อมูลรายละเอียดคลาสรถยนต์

ตารางที่ 2.8 อธิบายตัวอย่างคลาส และพร็อพเพอร์ตี้ของออนโทโลยีรถยนต์

เอนทิตี้	**ประเภทของข้อมูล (Type)**	**ความหมาย**
Car	คลาส (Class)	รถยนต์
Brand	คลาส (Class)	ยี่ห้อ
Honda	คลาส (Class)	รถยนต์ Honda
SubclassOf	พร็อพเพอรตี้ (Properties)	มีความสัมพันธ์แบบซับคลาส (Subclass)
Price	พร็อพเพอรตี้ (Properties)	ค่าคงที่แสดงถึงราคาของรถยนต์
hasPrice	พร็อพเพอรตี้ (Properties)	ระบุราคาของรถยนต์
hasBrand	พร็อพเพอรตี้ (Properties)	ระบุยี่ห้อของรถยนต์

2) ข้อมูลอินสแตนส์ (Instance Data) คือการอธิบายรายละเอียดของข้อมูลซึ่งใช้ข้อมูลเค้าร่างเป็นแม่แบบในการอธิบาย ดังภาพที่ 2.13 แสดงตัวอย่างข้อมูลอินสแตนส์อธิบายรายละเอียดรถยนต์



ภาพที่ 2.13 ตัวอย่างข้อมูลอินสแตนส์อธิบายรายละเอียดรถยนต์

การอธิบายข้อมูลเชิงความหมายที่นำไปใช้งานในโปรแกรมประยุกต์อาจจะเลือกอธิบายข้อมูลเค้าร่างเพียงอย่างเดียวก็ได้ ทั้งนี้ขึ้นกับความจำเป็นในการออกแบบ และใช้งานออนโทโลยีในโปรแกรมประยุกต์

สรุปได้ว่า ออนโทโลยีมีการกำหนดโครงสร้างความรู้ที่ชัดเจนของสิ่งที่เกี่ยวข้องกับแนวคิด (Concepts) เพื่อนำมาใช้ในการบรรยายหรือแสดงความคิดในขอบเขตความรู้เรื่องใดเรื่องหนึ่ง (Domain) สามารถแสดงให้เห็นถึงคุณสมบัติที่เกี่ยวข้องกับสิ่งนั้น ๆ ดังนั้นโครงสร้างออนโลยีจึงประกอบด้วย การกำหนดนิยามความหมายหรือการอธิบายเรื่องราวโดยใช้แนวคิด (Concept) คุณสมบัติ (Properties) ความสัมพันธ์ (Relationships) ตรรกะการแปลงความสัมพันธ์ (Axioms) และคำศัพท์หรือตัวอย่างข้อมูล (Instances)

2.7 ภาษาการสอบถาม

ภาษาสำหรับการสอบถามข้อมูลที่เก็บในรูปแบบอาร์ดีเอฟ มีดังนี้

- SPARQL

ภาษาสำหรับการสอบถามข้อมูลสปาเก้ิล (SPARQL) เป็นภาษาสืบค้นข้อมูลที่อยู่ในรูปแบบภาษาอาร์ดีเอฟ (Prud'hommeaux and Seaborne, 2007) ภาษาสปาเก้ิล มีความสามารถในการสืบค้นข้อมูลหลายแบบ เช่น แบบมีเงื่อนไข เป็นต้นสปาเก้ิล เป็นภาษาสำหรับดึงข้อมูลมาแสดง ซึ่งอยู่บนพื้นฐานของข้อมูลที่เป็นไปในรูปแบบของกราฟ หรือรูปแบบของ Triple ที่ประกอบด้วย Subject Predicate และ Object แนวความคิดของหลักการสืบค้นด้วยภาษาสปาเก้ิล ที่สามารถเปรียบเทียบกันได้กับคำสั่งภาษาเอสคิวแอล (Sequential Question Language: SQL) โดยแบ่งเป็นสองส่วนที่

สำคัญ ได้แก่ ส่วน SELECT และ WHERE ซึ่ง SELECT จะไปอธิบายตัวแปร (โดยใช้ prefix “?” และ “$”) ซึ่งไปปรากฏบนผลลัพธ์ และWHERE จะเป็นเงื่อนไขดังแสดงดังตัวอย่าง

รูปแบบคำสั่งสืบค้นด้วยสปาเกิ้ล

```
SELECT ?subject ?object
WHERE { ?subject rdfs:subClassOf ?object }
```

```
Select ?subject ? object
WHERE {?subject :flowerName”มะลิ”.
?subject : hasUsefully ?object}
```

```
Select DISTINCT?x?z
WHERE {?x :Praewa_Silk_Cloth_Name"ผ้าผืนหรือผ้าชิ้น 5 เมตร".
?x :is_Praewa_Silk_Type_Patterns?z.
{?y :Praewa_Silk_Type_Patterns_Name?s.
{?w:Praewa_Silk_Patterns_Name?a.
FILTER regex(?a, "บังขาเข")}}}
```

สรุปได้ว่าสปาเกิ้ล เป็นภาษาสำหรับการสืบค้นข้อมูลจากออนโทโลยีอย่างสื่อความหมาย ซึ่งจะมีการเข้าถึงข้อมูลโดยอาศัยโครงสร้าง Triple (Subject, Predicate, Object) แนวความคิดของหลักการสืบค้นด้วยภาษาสปาเกิ้ล โดยแบ่งเป็นสองส่วนหลักๆ คือ ส่วน SELECT และ WHERE ซึ่ง SELECT จะไปอธิบายตัวแปรที่ใช้ prefix “ ? ” ซึ่งเป็นตัวแปรที่เก็บค่าผลลัพธ์ และ WHERE จะเป็นเงื่อนไขการสืบค้นข้อมูล ซึ่งมีการนิยามการสืบค้นออนโทโลยีด้วยภาษาสปาเกิ้ล ดังนี้

Sparql $\longrightarrow$ SELECT varlist WHERE (bgp)

โดยที่ varlist = (v1,v2,...,vn) varlist $\subseteq$ var(bgp)

bgp คือ Basic Graph Pattern

อาร์ดีคิวแอล (RDF Data Query Language: RDQL) (Oldakowski, 2004; Seaborne, 2004) เป็นภาษาในการสืบค้นข้อมูลบนโครงสร้างข้อมูลอาร์ดีเอฟ โดยรูปแบบของภาษาอาร์ดีคิวแอล

มีลักษณะที่คล้ายคลึงกับภาษาในการสอบถามข้อมูลแบบโครงสร้างภาษาเอสคิวแอลของฐานข้อมูลเชิงสัมพันธ์ ต่างกันตรงที่อ้างอิงข้อมูลทุกครั้งจะต้องอ้างถึง URI ด้วย

ภาษาอาร์ดีคิวแอล สามารถนำไปใช้ได้หลายภาษาไม่ว่าจะเป็นภาษา PHP หรือภาษา Perl ถึงแม้ว่า Syntax จะต่างกันแต่ก็มีหลักการเหมือนกัน โดยรูปแบบของการกำหนดตัวแปรในภาษาอาร์ดีคิวแอล จะใช้เครื่องหมาย ? นำหน้าชื่อตัวแปรเสมอ เช่น ? <identifier> และต้องมีเครื่องหมาย comma (,) คั่นระหว่างตัวแปรแต่ละตัว ภาษาอาร์ดีคิวแอล มีรูปแบบดังต่อไปนี้

```
SELECT variables  listing  FROM  rdf documents WHERE patterns AND
filter expressions  USING prefix declaration
```

ภาษาอาร์ดีคิวแอล มี Clause ซึ่งประกอบไปด้วย

1. SELECT clause เป็นการระบุรายการของตัวแปรที่ต้องการจากเงื่อนไข WHERE ซึ่งตัวแปรในภาษาอาร์ดีคิวแอล จะขึ้นต้นด้วยเครื่องหมาย ? และถ้ามีหลายตัวจะคั่นด้วยเครื่องหมาย comma (,) เช่น SELECT ?name ?email, ?age, ?tel_number หรือถ้าต้องการตัวแปรทั้งหมด จะใช้เครื่องหมาย * เช่น SELECT *

2. FROM clause ใช้ในการระบุ Path หรือ URL ของเอกสาร RDF ที่ต้องการสืบค้น ข้อมูลซึ่งในกรณีที่จะส่ง RDQL Statement ด้วยโปรแกรมไม่ต้องระบุ Path ก็ได้

3. WHERE clause เป็นการระบุรายการในรูปแบบ Triple ที่ต้องการผลลัพธ์ คือ (Subject, Predicate, Object) เช่น WHERE (?Resource, <http://www.w3.org/2001/vcard-rdf/3.0/EMAIL>,"radol@gmx.de") หมายถึง ต้องการค้นหา ?Resource ที่มีอีเมล์เป็น radol@gmx.de

4. USING ช่วยในการอธิบาย URI ที่ใช้ให้อยู่ในรูปแบบชื่อสั้น ๆ เพื่อให้ง่ายต่อการอ่านคำสั่งและการพิมพ์ เช่น WHERE (?Resource, vCard:EMAIL, "radol@gmx.de") USING vCard FOR <http://www.w3.org/2001/vcard-rdf/3.0/> หมายถึงใช้ vCard เป็นชื่อย่อของ URI http://www.w3.org/2001/vcard-rdf/3.0/

การสืบค้นของภาษาอาร์ดีคิวแอล นั้นใช้หลักในการสืบค้น คือ อาศัยการเชื่อมโยงของ Triple จำนวนมากที่เชื่อมต่อกันอยู่ในกราฟ โดยภาษาอาร์ดีคิวแอล จะมีรูปแบบในการสืบค้นโดยจะอ้างอิงไปยัง Triple แต่ละ Triple จนกว่าจะสามารถหาการเชื่อมโยงไปจนถึง Triple ที่มีค่าของข้อมูลที่เป็นผลลัพธ์ที่ต้องการ

รูปแบบในการสืบค้นของภาษาอาร์ดีคิวแอล นั้นมีรูปแบบคล้ายกับแบบสอบถามพื้นฐานของภาษาเอสคิวแอลบางส่วน เช่น SELECT WHERE ซึ่งมีรูปแบบดังต่อไปนี้

```
SELECT ?A, ?B
WHERE (Subject1,Predicate1,Object1 or Literal1),
(Subject2,Predicate2,Object2 or Literal2),(...)
USING ชื่อ Prefix FOR <URI ของ NameSpace>
```

จากรูปแบบในการสืบค้นของภาษาอาร์ดีคิวแอล ข้างต้น จะพบว่าตัวแปร ?A และ ?B เป็นตัวแปรที่ใช้แทนผลลัพธ์ที่ได้จากการสืบค้นข้อมูลจากเอกสารอาร์ดีเอฟ จากเงื่อนไขตามที่กำหนด และในส่วน (Subject1,Predicate1,Object1 or Literal1) แสดงถึงลำดับในการอ้างอิงถึงในแต่ละส่วนของ Triple ตามลำดับ โดยเส้นทางการสืบค้นตามภาษาอาร์ดีคิวแอล นั้น จะสืบค้นไปตามลูกศรโดยจะอ้างอิงไปยัง Triple แต่ละ Triple จนกว่าจะเชื่อมโยงไปถึง Triple ที่มีค่าของข้อมูลที่เป็นผลลัพธ์ แต่ลำดับในการอ้างอิงถึงแต่ละส่วนประกอบใน Triple จะอ้างอิงไปตามลำดับลูกศรที่เชื่อมระหว่างโหนดใน Triple

- **การวัดประสิทธิภาพการค้นหา**

ระบบการสืบค้นที่ดีจะต้องเป็นระบบที่สามารถดึงเอาข้อสนเทศที่เกี่ยวพันธ์ อันเป็นความต้องการและเป็นประโยชน์ได้อย่างรวดเร็ว ถูกต้องแม่นยำ ครบถ้วนสมบูรณ์ ดังนั้นการประเมินหรือการวัดประสิทธิภาพการค้นหา จึงมีความสำคัญเป็นอย่างมากสำหรับการค้นหา (ศุภชัย ตั้งวงศ์ศานต์, 2553) มีดังนี้

1) ค่าการจำได้ (Recall)

ค่าการจำได้ (Recall) ค่าระลึก หรือค่าความครบถ้วน ที่แสดงให้เห็นว่าเมื่อระบบได้ทำการดึงคำตอบออกมาแล้วมีความถูกต้องเพียงใด โดยสามารถคำนวณได้จากสูตรดังสมการ 2.1 ดังนี้

$$\boldsymbol{Recall: r = \frac{|Ra|}{|R|}} \qquad (2.1)$$

ค่าการจำได้เป็นการวัดความสามารถของระบบการสืบค้นในการดึงเอกสารทั้งหมดหรือข้อมูลที่เกี่ยวพันธ์กับคำสอบถาม หากค่า Recall เป็น 1 หรือ Ra = R หมายถึง เอกสารหรือข้อมูลที่ดึงออกมาทั้งหมดสอดคล้องกับคำสอบถามนั้น แต่ถ้าค่า Recall เป็น 0 หรือ Ra = 0 หมายถึง ไม่มีเอกสารหรือข้อมูลที่ดึงออกมาจะสอดคล้องกับคำถาม

2) ค่าความแม่นยำ (Precision)

ค่าความแม่นยำเป็นค่าที่จะแสดงให้เห็นว่าระบบที่พัฒนาขึ้นมามีความแม่นยำเพียงใด โดยสามารถคำนวณได้จากสูตรดังสามการ 2.2 ดังนี้

$$\text{Precision: p} = \frac{|\text{Ra}|}{|\text{A}|} \qquad (2.2)$$

ค่า Precision เป็นการวัดความสามารถของระบบการสืบค้นในการดึงเอกสารหรือข้อมูลที่เป็นคำตอบได้สอดคล้องตรงประเด็นกับคำสอบถามให้ได้มากที่สุด หากค่า Precision เป็น 1 หรือ Ra = A หมายถึง คำตอบของเอกสารหรือข้อมูลทุกรายการมีความเกี่ยวพันธ์กัน แต่ถ้า Precision เป็น 0 หรือ Ra = 0 หมายถึง ไม่มีคำตอบของเอกสารหรือข้อมูลที่มีความเกี่ยวพันธ์กัน

สมมุติว่า กลุ่มเอกสารหรือข้อมูลทั้งหมดเมื่อได้ทำการสืบค้นด้วยกฎเกณฑ์ใด ๆ เอกสารที่สอดคล้องกับคำสอบถาม หรือคำค้น จะแสดงได้ในลักษณะตารางแบบ 2 ทาง ดังตารางที่ 2.10 โดยในแนวตั้งคือ ส่วนเอกสารที่ถูกดึงออกมาเป็นคำตอบ (Retrieved) และส่วนเอกสารที่ไม่ได้เป็นคำตอบ (Not Retrieved) ในแนวนอนจะเป็นส่วนเอกสารที่มีคุณสมบัติเกี่ยวพันธ์กับคำสอบถามนั้น (Relevant) และส่วนที่เหลือที่ไม่ได้เกี่ยวข้อง (Not Relevant)

ตารางที่ 2.9 แสดงผลการสืบค้น

	Retrieved	NOT Retrieved	Total
Retrieved	Ra	R-Ra	R
NOT Retrieved	A-Ra	N-R-A+Ra	N-R
Total	A	N-A	N

ให้ N เป็นเซทของเอกสารหรือข้อมูลทั้งหมด

R เป็นเซทของเอกสารหรือข้อมูลทั้งหมดที่มีคุณสมบัติที่เกี่ยวข้องกับคำสอบถาม

A เป็นเซทของเอกสารหรือข้อมูลทั้งหมดที่เป็นคำตอบ

Ra เป็นเซทของเอกสารหรือข้อมูลที่เป็นคำตอบ และเกี่ยวพันธ์กับคำสอบถาม

การวัดความสามารถในการสืบค้นเอกสารหรือข้อมูลที่เกี่ยวพันธ์กันออกมาได้เป็นสิ่งสำคัญ ดังนั้น ทั้งค่า Recall และ Precision จึงเป็นหลักในการประเมินผลประสิทธิภาพของการสืบค้น

3) อัตราการรู้จำ (F-measure) เป็นค่าเฉลี่ยที่ให้ความสำคัญกับความแม่นยำและความครบถ้วนเท่า ๆ กัน โดยเป็นการเปลี่ยนค่า Recall และค่า Precision มารวมเป็นหนึ่งเดียว

งานวิจัยนี้ได้ประยุกต์ใช้ทฤษฎีของ (Miao et al, 2009) สำหรับเป็นเกณฑ์ในการประเมินประสิทธิภาพการสืบค้นด้วยอัตราการรู้จำ (F-Measure) ซึ่งเกิดจากการรวมเอาค่าความแม่นยำ (Precision) และค่าการจำได้ (Recall) มาใช้ในการคำนวณ ซึ่งคำนวณได้จากสูตรดังสมการ 2.3 ดังนี้

$$\text{F-Measure} = (2 \text{ x Precision x Recall})/(\text{Precision}+\text{Recall}) \quad (2.3)$$

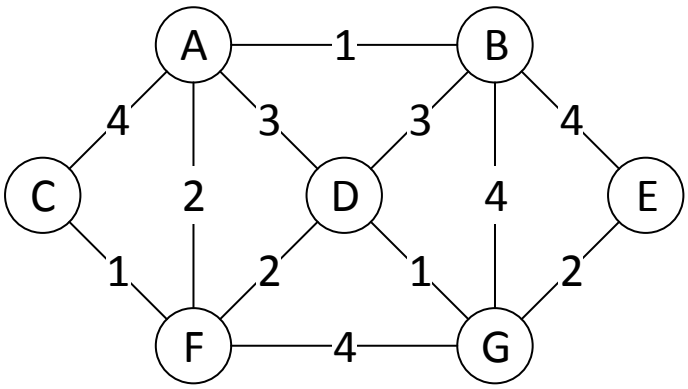
2.8 การหาเส้นทางที่สั้นที่สุด

แมคคอนเนลล์ (McConnell, 2001) ได้อธิบายไว้ว่า อัลกอริทึมเส้นทางที่สั้นที่สุดจะค้นหาชุดของเส้นเชื่อม (edge) ระหว่างสองโหนด ที่มีผลรวมน้ำหนักน้อยที่สุด อาจมีความคล้ายกับอัลกอริทึม minimum spanning tree ซึ่งเลือกเส้นทางที่มีน้ำหนักน้อยสุดแล้วหาจุดเชื่อมต่อไป แต่มีความแตกต่างกันเนื่องจากการหาเส้นทางสั้นที่สุดจะค้นหาจากทุกเส้นทางที่เป็นไปได้

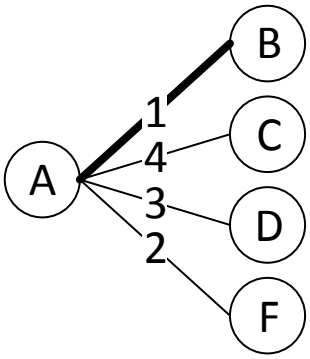
Dijkstra's Algorithm

```
เลือกโหนดเริ่มต้น
สร้าง fringe เริ่มต้นจากโหนดที่เชื่อมต่อกับโหนดเริ่มต้น
WHILE ขณะที่ยังไม่พบโหนดปลายทาง
        เลือกโหนด fringe กับเส้นทางสั้นที่สุดสู่โหนดเริ่มต้น
        เพิ่มโหนดนั้นและเส้นเชื่อมสู่ต้นไม้
        ปรับปรุง fringe โดย :
                เพิ่มโหนดสู่ fringe ที่เชื่อมต่อสุ่โหนดูใหม่
                For each node in the fringe do
                FOR ทุกโหนดใน fringe
                        ปรับปรุงเส้นเชื่อมสู่หนึ่งการเชื่อมต่อสู่ต้นไม้บนเส้นทางสั้นที่สุดสู่โหนดเริ่มต้น
                 END FOR
END WHILE
```

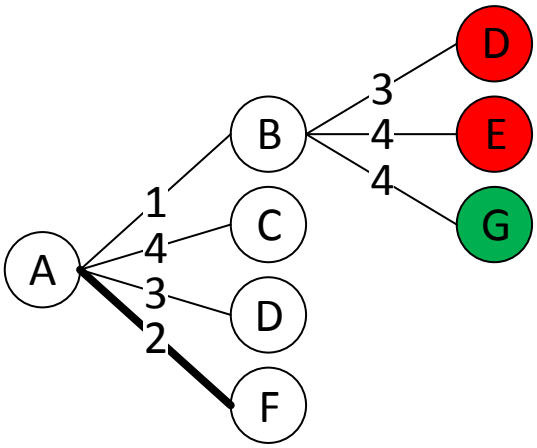
ตัวอย่างการหาค่าเส้นทางสั้นที่สุด จากโหนด A ไปยังโหนด G ตามโครงสร้างต่อไปนี้



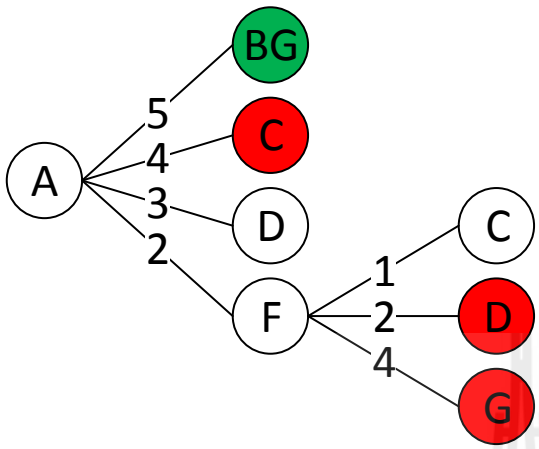
เริ่มจากการกำหนดโหนด A เป็นโหนดเริ่มต้น จากนั้นหาเส้นเชื่อมจากโหนดต่างๆ ถึงโหนด A เลือกเส้นเชื่อมที่มีค่าน้ำหนักต่ำสุดในที่นี้คือโหนด B มีค่าน้ำหนักเป็น 1



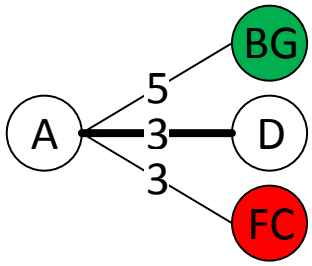
หาเส้นเชื่อมจากโหนดต่างๆ ถึงโหนด B หากเส้นเชื่อมมีโหนดร่วมกับเส้นทางอื่นให้ดูค่าน้ำหนักรวมจากจุดเริ่มต้นแล้วเลือกเส้นทางที่มีค่าน้ำหนักรวมมีค่าต่ำสุดไว้แล้วตัดเส้นทางอื่นทิ้ง ดังเส้นทาง ABD กับ เส้นทาง AD ซึ่งเส้นทาง ABD ถูกตัดทิ้งเนื่องจากมีค่าน้ำหนักร่วมสูงกว่า กรณีพบว่ามีเส้นทางจากโหนดต้นทางถึงปลายทางแล้ว เส้นทางอื่นๆที่มีค่าน้ำหนักรวมมีค่าต่ำสุดสูงกว่าเส้นทางนั้นจะถูกตัดทิ้งด้วย ดังเส้นทาง ABE จากนั้นหาเส้นเชื่อมจากโหนดต่างๆ ถึงโหนด A เลือกเส้นเชื่อมที่มีค่าน้ำหนักต่ำสุดในที่นี้คือโหนด F มีค่าน้ำหนักเป็น 2



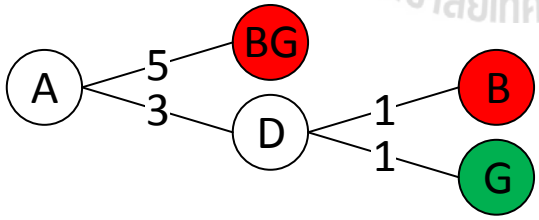
หาเส้นเชื่อมจากโหนดต่างๆ ถึงโหนด F เส้นทาง AC ถูกตัดทิ้งเนื่องจากมีค่าน้ำหนักร่วมสูงกว่าเส้นทาง AFC เส้นทาง AFD ถูกตัดทิ้งเนื่องจากมีค่าน้ำหนักร่วมสูงกว่าเส้นทาง AD และ เส้นทาง AFG ถูกตัดทิ้งเนื่องจากมีค่าน้ำหนักร่วมสูงกว่าเส้นทาง ABG



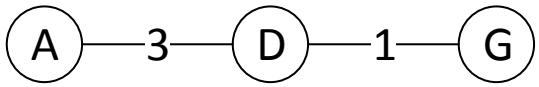
เส้นทาง AFC ถูกตัดทิ้งเนื่องจากไม่มีโหนดเชื่อมต่อ จากนั้นหาเส้นเชื่อมจากโหนดต่างๆ ถึงโหนด A เลือกเส้นเชื่อมที่มีค่าน้ำหนักต่ำสุดในที่นี้คือโหนด D มีค่าน้ำหนักเป็น 3



เส้นทาง ADB ถูกตัดทิ้งเนื่องจากมีค่าน้ำหนักร่วมสูงกว่าเส้นทาง AB และ เส้นทาง ABG ถูกตัดทิ้งเนื่องจากมีค่าน้ำหนักร่วมสูงกว่าเส้นทาง ADG



ผลลัพธ์สุดท้ายคือเส้นทางสั้นที่สุด คือ ADG มีค่าน้ำหนักร่วมเท่ากับ 4



สำหรับงานวิจัยนี้ ได้ประยุกต์ใช้ทฤษฎีการหาเส้นทางที่สั้นที่สุดสำหรับการค้นหาเอกสารความรู้ทางการเกษตรจากฐานความรู้ที่จัดทำขึ้นในรูปแบบออนโทโลยี เนื่องจากว่าในฐานความรู้ที่จัดทำขึ้นจะมีการกำหนดเอกสารสำหรับคำค้นไว้ และคำค้นต่าง ๆ เหล่านั้นจะถูกจัดความสัมพันธ์ไว้ ดังนั้น เพื่อให้

สามารถค้นหาข้อมูลได้ตรงตามความต้องการของผู้ใช้มากที่สุด จึงได้นำทฤษฎีการหาเส้นทางที่สั้นที่สุดมาประยุกต์ใช้ในขั้นตอนวิธีการค้นหาเชิงความหมายที่ได้พัฒนาขึ้นมา

2.9 งานวิจัยที่เกี่ยวข้อง

กระบวนการจัดการฐานความรู้และค้นคืนความรู้ มีงานวิจัยที่เกี่ยวข้องเพื่อจัดการความรู้ให้เป็นระบบเพื่อที่จะได้ค้นคืนความรู้ต่าง ๆ ได้อย่างรวดเร็ว และมีความถูกต้องตรงตามความต้องการของผู้ใช้ นอกจากนี้ยังมีงานวิจัยที่ได้นำเสนอเกี่ยวกับการแสดงผลการค้นคืนความรู้เหล่านั้น สำหรับงานวิจัยที่เกี่ยวข้อง แสดงตามงานวิจัยของผู้วิจัยแต่ละท่าน มีดังนี้

นาถธิชา เกษรพันธ์ (2554) ได้ทำการวิจัย การพัฒนาออนโทโลยีสำหรับระบบให้คำแนะนำอาจารย์ผู้เชี่ยวชาญตามหัวข้องานวิจัยของนักศึกษา ซึ่งในงานวิจัยนี้ได้ทำการพัฒนาออนโทโลยีสำหรับระบบให้คำแนะนำอาจารย์ผู้เชี่ยวชาญตามหัวข้องานวิจัย เพื่อให้การแนะนำอาจารย์ผู้เชี่ยวชาญที่เหมาะสมและตรงตามวัตถุประสงค์ของงานวิจัย ลดปัญหาวิธีการวิจัยที่ไม่ชัดเจนส่งผลต่อประสิทธิภาพของผลงานวิจัยและลดความเสี่ยงต่อการวิจัยล้มเหลว โดยระบบจะมีส่วนในการเข้าถึงและแบ่งปันความรู้จากการเชื่อมโยงข้อมูลในฐานข้อมูลเข้ากับฐานความรู้ (Database – Ontology) ด้วยการใช้มาตรฐานเว็บเชิงความหมาย (Semantic Web Standards) เพื่อเพิ่มความอัตโนมัติของกระบวนการและเพิ่มประสิทธิภาพในการสืบค้นที่ให้คำแนะนำที่แม่นยำ จากผลการศึกษาพบว่าผู้ใช้งานระบบพึงพอใจกับการแนะนำผู้เชี่ยวชาญเพราะระบบมีช่องทางในการเข้าถึงข้อมูลภายในฐานข้อมูลหรือคลังข้อมูล โดยระบบสามารถแนะนำ รายชื่ออาจารย์ผู้เชี่ยวชาญ ตามข้อมูลหรือคำสำคัญที่ผู้ใช้ระบุและแสดงคุณสมบัติ ข้อมูลต่าง ๆ ของอาจารย์ผู้เชี่ยวชาญเพื่อให้ผู้ใช้สามารถนำข้อมูลไปตัดสินใจได้ต่อไป นอกจากนี้ระบบยังมีส่วนที่ใช้งานโดยผู้ให้ความรู้แก่ระบบ ซึ่งมีช่องทางป้อนข้อมูลหรือองค์ความรู้จากอาจารย์ผู้เชี่ยวชาญในแต่ละสาขาวิชา โดยระบบสามารถเพิ่ม ลบ แก้ไขข้อมูลหรือองค์ความรู้ของอาจารย์ผู้เชี่ยวชาญได้

เพ็ญพรรณ อัศวนพเกียรติ (2547) ได้ทำการวิจัย ระบบสำหรับสืบค้นและรวบรวมข้อมูลทางชีววิทยาโดยใช้ชีวออนโทโลยี ซึ่งวิทยานิพนธ์นี้ได้พัฒนาระบบสำหรับรวบรวมและสืบค้นข้อมูลทางชีววิทยา ระบบประกอบด้วย 3 ส่วน ได้แก่ การพัฒนาฐานความรู้ การสืบค้นและรวบรวมข้อมูล และการนำเสนอรูปแบบการสืบค้นต่อผู้ใช้ ในส่วนการพัฒนาฐานความรู้งานวิทยานิพนธ์นี้ได้นำองค์ความรู้ที่มีอยู่มาพัฒนา ในส่วนการรวมข้อมูลงานนี้ได้ใช้เทคนิคปัญญาประดิษฐ์ด้วยการพิจารณาปัจจัยที่มีต่อการสืบค้นเพื่อเพิ่มประสิทธิภาพในการสืบค้น ในรูปแบบการสืบค้นงานวิจัยนี้นำเสนอการนำข้อดีของภาษาสืบค้นและการใช้ภาพเพื่อสื่อความเข้าใจมารวมกันเพื่อสนับสนุนข้อมูลคำถามที่ซับซ้อน ส่วนหลักของวิทยานิพนธ์นี้ประกอบด้วย การพัฒนาองค์ความรู้ทางชีววิทยา เทคนิคใหม่ในการรวมข้อมูล และการนำภาษาสืบค้นและรูปภาพเพื่อสื่อความเข้าใจมารวมกัน เมื่อเปรียบเทียบองค์ความรู้ที่มีอยู่ และงานวิจัยนี้พบว่าองค์ความรู้ของระบบนี้มีความเหมาะสมมากกว่าองค์ความรู้อื่น ในการรวมภาษาสืบค้น

และการใช้ภาพเพื่อสื่อความหมายระบบนี้สามารถสืบค้นคำถามที่ซับซ้อนได้ และใช้งานง่ายเหมาะสำหรับผู้เริ่มต้นและผู้เชี่ยวชาญ

กานดา ศรอินทร์ (2548) ได้ทำการวิจัยระบบสืบค้นข้อมูลเฉพาะเจาะจงโดยใช้คำเฉพาะและออนโทโลยี ซึ่งงานวิจัยนี้ได้สร้างระบบสืบค้นข้อมูลเฉพาะเจาะจงโดยใช้คำเฉพาะและออนโทโลยี ภายใต้โดเมนที่ถูกจำกัดไว้อย่างชัดเจน ขั้นตอนในการออกแบบระบบประกอบด้วย 1) การบรรยายโครงสร้างความสัมพันธ์ของระบบผ่านภาษาโอดัลเบิ้ลยูแอล 2) การบรรยายฐานข้อมูลที่ใช้สืบค้นผ่านภาษา XML 3) แบบจำลองเวกเตอร์สเปซโมเดลถูกนำมาใช้เพื่อแสดงผลการสืบค้นโดยจัดเรียนผลลัพธ์ตามลำดับชั้นของความถี่หมวดเอกสาร 4) เทคนิคการค้นหาแบบบูลีนถูกนำมาใช้เพื่อกำหนดรูปแบบในการสืบค้นของผู้ใช้งาน สำหรับในงานวิจัยนี้ได้เลือกโดเมนของระบบจัดเก็บเอกสารอิเล็กทรอนิกส์มหาวิทยาลัยขอนแก่น มาสร้างระบบสืบค้นข้อมูลเฉพาะเจาะจงโดยใช้คำเฉพาะและออนโทโลยีตามขั้นตอนและวิธีที่เสนอไว้ข้างต้น เมื่อนำผลลัพธ์ที่ได้จากระบบสืบค้นข้อมูลดังกล่างมาเปรียบเทียบกับระบบสืบค้นแบบไดเรคทอรี พบว่าระบบสืบค้นข้อมูลแบบเฉพาะเจาะจงโดยใช้คำเฉพาะและออนโทโลยีมีการเข้าถึงข้อมูลที่เร็วกว่าเนื่องจากใช้การอ้างอิงข้อมูลแบบระบบเวกเตอร์ที่สามารถเข้าถึงอิลิเมนต์ข้อมูลได้โดยตรง รวมทั้งมีการจัดลำดับชั้นเอกสารอัตโนมัติตามลำดับความถี่ของหมวดเอกสารไปยังผู้ใช้ในขณะที่ระบบสืบค้นข้อมูลแบบไดเรคทอรีอาศัยการสืบค้นข้อมูลในระบบฐานข้อมูลเชิงสัมพันธ์โดยใช้ภาษาเอสคิวแอล ซึ่งใช้เวลามากในการค้นหาคำในประโยคที่มีขนาดใหญ่ รวมถึงผู้ใช้ต้องเลือกหมวดเอกสารเพื่อให้ระบบสามารถแสดงผลลัพธ์แบบเฉพาะเจาะจงตามที่ระบุไว้ ในขณะที่ระบบสืบค้นข้อมูลเฉพาะเจาะจงโดยใช้คำเฉพาะและออนโทโลยีมีการแยกย่อยคำจากประโยคขนาดใหญ่มาจัดเก็บและอ้างอิงในระบบเวกเตอร์ การอ้างอิงดังกล่าวทำให้สามารถเข้าถึงอิลิเมนต์ข้อมูลที่ใช้แสดงผลได้โดยตรงผ่านกระบวนการ SAX ซึ่งมีการทำงานและเข้าถึงข้อมูลแบบเรียงลำดับ ประโยชน์ที่ได้ทำให้สามารถค้นหาและเข้าถึงข้อมูลได้อย่างรวดเร็ว รวมทั้งสามารถแสดงผลตามความถี่ของหมวดเอกสารอัตโนมัติ อีกทั้งผู้ใช้ยังสามารถกำหนดรูปแบบในการค้นหาได้ตามความต้องการโดยใช้เทคนิคการค้นหาแบบบูลีนร่วมกับคำเฉพาะจากผู้ใช้

จารุณี แซ่หลี (2549) ได้ทำการวิจัย การค้นหาคำสำคัญในฐานข้อมูลโดยใช้การค้นข้อมูลฐานข้อมูล ซึ่งงานวิจัยนี้ได้เสนอวิธีการค้นหาคำสำคัญในฐานข้อมูล ซึ่งผู้ใช้สามารถใช้คำสำคัญหรือกลุ่มคำสำคัญที่ตรงกับค่าข้อมูล ข้อมูลฐานข้อมูล และคำพ้องความหมายในการค้นหาได้ โดยอาศัยแบบจำลองข้อมูลที่ประกอบด้วยข้อมูล 3 ระดับ คือ โครงสร้างฐานข้อมูล คำพ้องความหมาย และเคส เรียกว่า การค้นหาคำสำคัญในฐานข้อมูลโดยใช้การค้นข้อมูลฐานข้อมูล เมื่อทำการเปรียบเทียบผลลัพธ์ที่ดีที่สุดที่ได้จากการค้นหาด้วยระบบการค้นหาคำสำคัญในฐานข้อมูลระบบต่างๆ พบว่า การค้นหาคำสำคัญในฐานข้อมูลโดยใช้การค้นข้อมูลฐานข้อมูลให้ผลลัพธ์ที่ถูกต้องแม่นยำกว่าระบบที่ใช้การค้นหาคำสำคัญที่ตรงกับค่าข้อมูลเพียงอย่างเดียว โดยสามารถกำหนดผลลัพธ์ได้จากการเลือกระเบียนข้อมูล การ

เลือกแอตทริบิว และการเลือกเอนทิตี้ ในขณะที่ระที่ใช้การค้นหาคำสำคัญที่ตรงกับค่าข้อมูลเพียงอย่างเดียวสามารถกำหนดผลลัพธ์ได้จากการเลือกระเบียนข้อมูลเท่านั้น

นคร โคตรโสภา (2549) ได้ทำการวิจัย วิธีการสืบค้นเอกสารเชิงความหมายบนข้อมูลเอ็กซ์เอ็มแอล ซึ่งงานวิจัยนี้ได้นำเสนอวิธีการสืบค้นเอกสารเชิงความหมายบนข้อมูลเอ็กซ์เอ็มแอล โดยแบ่งการทำงานเป็น 3 ขั้นตอน คือ 1) การทำดัชนีคำหลัก (Indexing) เป็นขั้นตอนในการแบ่งเอกสารเอ็กซ์เอ็มแอล ออกเป็นกลุ่ม เพื่อสร้างความสัมพันธ์ของเอกสารตามคำหลัก โดยกำหนดระดับความสำคัญและคำนวณน้ำหนักความหมายของคำหลักเพื่อหาระดับความสัมพันธ์ระหว่างกลุ่มเอกสารกับคำหลัก 2) การจัดลำดับผลลัพธ์ (Ranking) เป็นการใช้หลักการวัดระดับความคล้ายคลึงกันเชิงความหมายของข้อมูลคำหลักในเอกสารกับความหมายของข้อมูลที่สืบค้น เพื่อจัดเรียงลำดับเอกสารตามค่าระดับความคล้ายคลึงกันจากค่ามากไปหาน้อย 3) การสืบค้น (Searching) จากการทดลองตามแบบจำลองที่ออกแบบไว้พบว่าเอกสารที่มีคำหลักอยู่ในแท็กที่มีความสูงมากจะมีค่าน้ำหนักของคำหลักมาก แต่เอกสารที่มีความถี่ของคำหลักมากอาจจะมีค่าน้ำหนักของคำหลักต่ำ เพราะคำหลักอาจจะปรากฏในแท็กที่มีความสูงน้อย และเมื่อทดลองใส่คำค้น ผลลัพธ์ที่ได้จะมีการจัดลำดับตามค่าระดับความคล้ายคลึงกันจากค่ามากไปหาน้อย ทำให้ผู้ใช้งานได้เอกสารตามความหมายของข้อมูลที่สืบค้น

ชมพูนุช ภักติยานุวรรตน์ (2550) ได้ทำการวิจัย การบูรณาการออนโทโลยีอย่างมีความหมายจากแหล่งข้อมูลที่หลากหลาย ซึ่งงานวิจัยนี้มุ่งเน้นการแก้ปัญหาความหลากหลายในเชิงความหมายของข้อมูล และบูรณาการข้อมูลจากแหล่งข้อมูลที่หลากหลายโดยใช้คำอธิบายข้อมูลบนพื้นฐานของออนโทโลยีซึ่งได้ถูกพัฒนาให้อยู่ในรูปแบบของภาษาโอดับเบิ้ลยูแอล งานวิจัยนี้ยังสามารถประยุกต์กับกรณีศึกษา โดยได้แสดงแหล่งข้อมูลที่หลายหลายที่เกิดขึ้นในโดเมนมหาวิทยาลัย นอกจากนี้ยังได้นำเสนอกระบวนการสืบค้นข้อมูลโดยผ่านคำอธิบายข้อมูลนี้ด้วย ผลที่ได้รับคือการประยุกต์ใช้งานคำอธิบายข้อมูลจะช่วยในการแก้ปัญหาความขัดแย้งกันข้อมูลและการเก็บข้อมูลที่ซ้ำซ้อนกันในแหล่งข้อมูลที่หลากหลายได้ โดยมีกระบวนการวิจัยที่สำคัญอยู่ 3 กระบวนการคือ 1) ขั้นตอนการแปลงข้อมูลในรูปแบบฐานข้อมูลเชิงสัมพันธ์ให้อยู่ในรูปแบบออนโทโลยีโดยภาษาโอดับเบิ้ลยูแอล ใช้โปรแกรม Protégé เป็นเครื่องมือ 2) ขั้นตอนการเชื่อมโยงข้อมูลในรูปแบบที่เรียกว่าการ Alignment คือข้อมูลที่ได้ทำการเชื่อมโยงแล้วจะถูกเก็บไว้ที่ออนโทโลยีใดเพียงออนโทโลยีเดียว โดยใช้เครื่องมือ Altova Semantic Work 2006 3) ขั้นตอนการสืบค้นข้อมูลโดยใช้ภาษาสปาเก้ิล และ Reasoner กรณีแรกสืบค้นข้อมูลออนโทโลยีของมหาวิทยาลัยขอนแก่น หรือมหาวิทยาลัยเกษตรศาสตร์เพียงอย่างเดียวในการสืบค้นแต่ละครั้ง กรณีที่สองเป็นการสืบค้นคำที่เขียนเหมือนกันแต่มีความหมายต่างกัน และกรณีที่สามเป็นการสืบค้นคำที่เขียนต่างกันแต่มีความหมายเหมือนกัน

ประสิทธิ์ ประทุมรัตน์ (2550) ได้ทำการวิจัย การจัดเก็บและสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยี ซึ่งงานวิจัยนี้ได้นำเสนอแบบจำลองการจัดเก็บและสืบค้นรูปภาพเชิงความหมายโดยใช้ออน

โทโลยี โดยใช้ข้อมูลรูปภาพจดหมายเหตุของมหาวิทยาลัยขอนแก่นเป็นกรณีศึกษา โมเดลที่นำเสนอได้ทำการอธิบายข้อมูลรูปภาพด้วย Metadata ผ่านภาษาเชิงความหมาย อาร์ดีเอฟ และสร้างขบวนการสืบค้นข้อมูลรูปภาพให้อยู่ภายใต้โดเมนที่ชัดเจน โดยการสร้างออนโทโลยีรูปภาพจดหมายเหตุของ มหาวิทยาลัยขอนแก่นขึ้นมาเพื่อใช้สำหรับการค้นคืนรูปภาพที่ได้อธิบายข้อมูลตามโครงสร้างที่ได้ออกแบบไว้แล้ว จากการทดลองสืบค้นข้อมูลโดยใช้ข้อมูลตัวอย่างที่สร้างขึ้นมา ผลการวิจัยพบว่าตัวแบบที่นำเสนอสามารถแก้ปัญหาการสืบค้นข้อมูลบนเว็บที่มีลักษณะที่แตกต่างกัน 3 ประการคือ 1) คำที่เขียนเหมือนกันแต่มีความหมายต่างกัน 2) คำที่เขียนต่างกันแต่มีความหมายเหมือนกัน 3) ข้อมูลขาดการสร้างความสัมพันธ์อย่างชัดเจน

ภรศิษฐ์ กุหลาบทิพย์ (2550) ได้ทำการวิจัย การค้นหาข้อมูลบนเว็บเชิงความหมายโดยใช้ฐานความรู้ Wordnet ซึ่งผู้วิจัยได้สร้างเครื่องมือค้นหาข้อมูลบนเอกสารเว็บเพจเชิงความหมายโดยใช้ฐานความรู้ Wordnet โดยมุ่งเน้นไปที่ความหมายของข้อมูลหรือ Semantic Search เพื่อให้เกิดผลการสืบค้นข้อมูลที่ตรงกับความต้องการของผู้ใช้มากที่สุด มีขั้นตอนในการออกแบบระบบประกอบด้วย 1) สกัดข้อมูลสารสนเทศบนเอกสารเว็บเพจเพื่อเปรียบเทียบข้อมูลสารสนเทศเชิงความหมายกับฐานความรู้ Wordnet 2) จัดเก็บข้อมูลในรูปแบบโครงสร้างเอกสารเอ็กซ์เอ็มแอล 3) จัดทำดัชนีข้อมูลเอกสารเอ็กซ์เอ็มแอล 4) กำหนดค่าน้ำหนักของคำสำคัญเพื่อการจัดเรียงผลลัพธ์การสืบค้นให้กับผู้ใช้ ซึ่งผลการวิจัยพบว่าระบบสามารถค้นหาข้อมูลเชิงความหมายได้ดี โดยสามารถสืบค้นเจอกลุ่มคำอื่นๆ ที่มีความหมายอยู่ในกลุ่มเดียวกัน และจากการตรวจสอบความถูกต้องของข้อมูลพบว่า ผลลัพธ์ความถูกต้องของข้อมูลในเอกสารเว็บเพจที่สามารถสืบค้นเจอโดยวิธีการสืบค้นเชิงความหมาย สามารถสืบค้นได้ผลลัพธ์ของข้อมูลที่มีเนื้อหาและความหมายตรงกับคำสำคัญที่ใช้ในการสืบค้นมากกว่าผลลัพธ์ของความถูกต้องของข้อมูลที่ได้จากการสืบค้นโดยวิธีสืบค้นโดยใช้คำสำคัญ

วีรนันท์ มาลาศิลป์ (2550) ได้ทำการวิจัย การบูรณาการข้อมูลอย่างมีความหมายจากแหล่งข้อมูลที่หลากหลายโดยคำอธิบายข้อมูลออนโทโลยี ซึ่งงานวิจัยนี้มุ่งเน้นการแก้ปัญหาความหลากหลายในเชิงความหมายของข้อมูลและการบูรณาการข้อมูลจากแหล่งข้อมูลที่หลากหลายโดยใช้คำอธิบายข้อมูลบนพื้นฐานของออนโทโลยี ซึ่งได้ถูกพัฒนาให้อยู่ในรูปแบบของภาษาโอดับเบิ้ลยูแอล งานวิจัยนี้ยังสามารถประยุกต์กับกรณีศึกษา โดยได้แสดงแหล่งข้อมูลที่หลากหลายที่เกิดขึ้นในโดเมนมหาวิทยาลัย อันได้แก่ ระบบงานบุคลากรและระบบข้อมูลวิจัย นอกจากนี้ยังได้นำเสนอกระบวนการสืบค้นข้อมูลโดยผ่านคำอธิบายข้อมูลนี้ด้วย ผลที่ได้รับคือการประยุกต์ใช้งานคำอธิบายข้อมูลจะช่วยในการแก้ปัญหาความขัดแย้งกับข้อมูลและเก็บข้อมูลที่ซ้ำซ้อนกันในแหล่งข้อมูลที่หลากหลายได้

บุญรัตน์ เผดิมรอด (2551) ได้ทำการวิจัยวิธีการใหม่สำหรับสืบค้นแบบหลายความสัมพันธ์บนองค์ความรู้ในฐานข้อมูลเชิงสัมพันธ์ ซึ่งงานวิจัยนี้เสนอแนวทางการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยีที่มีหลากหลายความสัมพันธ์เป็นฐานความรู้สำหรับการสืบค้นเชิงความหมาย

โดยจัดเก็บออนโทโลยีในฐานข้อมูลเชิงสัมพันธ์ รูปแบบความสัมพันธ์เชิงวัตถุที่พิจารณาถึงการอ้างอิงความสัมพันธ์ระหว่างโหนดในออนโทโลยีที่มีหลากหลายความสัมพันธ์ และพัฒนา ONT_RELATED โอเปอร์เรเตอร์ซึ่งจะแปลงคำสืบค้นของผู้ใช้ให้เป็นคำสืบค้นเชิงความหมายที่ใกล้เคียงกับคำสั่งของผู้ใช้มากที่สุด โดยโอเปอร์เรเตอร์นี้จะพิจารณากลุ่มคำในออนโทโลยีที่สอดคล้องกับคำสั่งสืบค้นตามความสัมพันธ์ที่กำหนด และนำกลุ่มคำที่ได้มาสืบค้นในฐานข้อมูล จากผลการทดลองเปรียบเทียบงานวิจัยการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยีที่พัฒนาโดยใช้คำสั่ง “START WITH” และ “CONNECT BY” ดังนั้นโอเปอร์เรเตอร์ที่เสนอในงานวิจัยสามารถสืบค้นข้อมูลได้อย่างมีประสิทธิภาพ และให้ผลการสืบค้นที่ครบถ้วน เมื่อพิจารณาถึงความครบถ้วนของผลการสืบค้น ONT_RELATED โอเปอร์เรเตอร์ที่เสนอให้ผลการสืบค้นที่ครบถ้วนกว่า การใช้คำสั่ง “START WITH” และ “CONNECT BY” เมื่อความสัมพันธ์ระหว่างคำที่ต้องการสืบค้นกับโหนดลูกในออนโทโลยีมีความสัมพันธ์ที่หลากหลาย โดยพิจารณาจากการประมาณค่าการสูญหายของผลการสืบค้น เนื่องจากงานวิจัยนี้จะพิจารณาถึงการอ้างอิงความสัมพันธ์ระหว่างโหนดในออนโทโลยีที่มีหลายความสัมพันธ์ โดยกำหนดความสัมพันธ์ตามค่าความเหมือนเชิงความหมาย และค่าความสำคัญที่กำหนดโดยผู้เชี่ยวชาญ นอกจากนี้ประโยชน์ที่ได้รับจากการจัดเก็บออนโทโลยีในฐานข้อมูลเชิงสัมพันธ์ คือการจัดการออนโทโลยี เช่น แก้ไข ลบ และการเข้าถึงออนโทโลยีสามารถทำได้ง่ายโดยใช้คำสั่งเอสคิวแอล

ปฏิคม ทองจริง (2552) ได้ทำการวิจัย ออนโทโลยีสำหรับการรวมข้อมูลเชิงความหมายของความรู้ด้านสมุนไพรไทย ซึ่งงานวิจัยนี้นำเสนอออนโทโลยีทางด้านสมุนไพรไทยที่ถูกจัดเก็บในรูปแบบของภาษาโอดับเบิ้ลยูแอลเพื่อนำมาใช้ในการรวมข้อมูลเชิงความหมายอย่างเป็นขั้นตอน นำความสัมพันธ์ของคำและความหมายของคำที่เกี่ยวข้องกับสมุนไพรไทยมาใช้ในการค้นคืนข้อมูล โดยข้อมูลสมุนไพรไทยจะถูกส่งผ่านสถาปัตยกรรม Mediator-Wrapper ที่ถูกกำหนดขึ้นเพื่อดึงข้อมูลจากแหล่งเว็บไซต์ต่าง ๆ ที่น่าเชื่อถือทางด้านสมุนไพร และรวมรวมข้อมูลเหล่านี้เพื่อนำเสนอรายละเอียดแก่ผู้ใช้งาน ทำให้เข้าใจความหมาย สรรพคุณ คุณสมบัติ ในขอบเขตด้านสมุนไพรไทยที่ครอบคลุมในทุกหมวดหมู่ของสมุนไพรตามเหตุผลที่ผู้ใช้ต้องการ อีกทั้งยังมีความสะดวกในการเพิ่มเว็บไซต์ที่เป็นแหล่งข้อมูลเพิ่มเติมเข้ามาจากการเรียกใช้ฟังก์ชั่นที่พัฒนาขึ้น และสามารถถูกเลือกใช้เป็นแหล่งข้อมูลใหม่ได้ โดยข้อมูลที่ได้จากเว็บไซต์ที่เพิ่มเข้ามาจะสามารถเพิ่มเนื้อหาของหมวดข้อมูลที่ขาดหายไปของบางสมุนไพรที่ไม่ค้นพบจากค่าเว็บไซต์เริ่มต้นที่กำหนดไว้ งานวิจัยนี้ทำให้การค้นคืนข้อมูลในขอบเขตความรู้สมุนไพรไทยมีประสิทธิภาพและสะดวกต่อการใช้งาน และยังนำไปซึ่งการเพิ่มประสิทธิภาพกับงานวิจัยที่เกี่ยวกับการใช้สมุนไพรไทยทางการแพทย์ต่อไป

มนัสนันท์ ปัญญามี (2552) ได้ทำการวิจัย การจัดกลุ่มผลการสืบค้นเอกสารบนเว็บเพื่อการสืบค้นเชิงความหมาย ซึ่ง งานวิจัยนี้ได้เสนอแนวคิดการแสดงความหมายของเว็บเพจ โดยการจัดกลุ่มเว็บเพจที่เป็นผลการสืบค้นที่ผู้ใช้เลือกคลิกดูเว็บเพจจากรายการผลการสืบค้นทั้งหมดที่แสดงโดย Search Engine เพราะการที่ Search Engine แสดงหัวข้อของเว็บเพจหรือข้อความที่ตัดมาบางส่วนจะเป็น

ข้อมูลเบื้องต้นให้ผู้ใช้เลือกเข้าไปดูหน้าเว็บที่คาดว่าจะมีข้อมูลที่ต้องการ การเก็บข้อมูลคำสืบค้นและเว็บเพจที่ผู้ใช้เลือกทำให้พบว่าสถิติของการเลือกใช้คำสืบค้นและเว็บเพจที่ถูกเลือกดูมีความสัมพันธ์กัน สามารถนำความสัมพันธ์นี้มาอธิบายได้ว่าเว็บเพจที่ผู้ใช้นิยมเลือกเข้าไปดูข้อมูลเป็นเว็บเพจที่มีความหมายตรงกับคำสืบค้นที่ผู้ใช้ป้อน และงานวิจัยนี้ยังได้หาความสัมพันธ์ในเชิงความหมายของคำสืบค้นที่ผู้ใช้ป้อนโดยใช้ฐานข้อมูล WordNet เพื่อจัดกลุ่มคำที่มีความหมายเหมือนกัน หลังจากที่ได้ทำการทดลองจัดกลุ่มผลการสืบค้นแล้วพบว่าสามารถวัดความแม่นยำ (Precision) ได้ 91.70% ค่าการจำได้ (Recall) ได้ 65.40% และค่าอัตราการรู้จำได้ 76.30% ได้กลุ่มของเว็บเพจที่มีความหมายโดยรวมเหมือนกันหรือเป็นเว็บเพจที่กล่าวถึงเรื่องเดียวกัน สามารถสืบค้นเอกสารที่มีความเกี่ยวข้องหรือเชื่อมโยงกันในเชิงความหมายถึงแม้เอกสารนั้นจะไม่มีคำค้นที่ผู้ใช้ป้อนเข้าไปปรากฏอยู่ และสามารถลดจำนวนผลการสืบค้นที่ไม่มีข้อมูลเกี่ยวข้องกับหัวข้อในการสืบค้นน้อยลงด้วย

สมมณี ลูซะวงษ์ (2553) ได้ทำการวิจัย การเข้าถึงฐานข้อมูลบนพื้นฐานของออนโทโลยีด้วยวิธีการปรับเปลี่ยนคำสั่งการสืบค้น ซึ่งงานวิจัยนี้ได้แสดงเทคนิคการปรับเปลี่ยนคำสั่งการสืบค้นสำหรับการเข้าถึงฐานข้อมูลโดยการใช้วิธีการออนโทโลยีเข้ามาช่วย จุดประสงค์หลักของงานวิจัยคือการแก้ปัญหาการเข้าถึงฐานข้อมูลที่ใช้ภาษาเอสคิวแอลที่เป็นเหมือนกับการสืบค้นด้วยคำเฉพาะ ซึ่งผู้ใช้จะต้องรู้จักชื่อตารางและคอลัมน์ของฐานข้อมูลนั้น ๆ และเพื่อที่จะให้ได้ผลลัพท์ของการสืบค้นมีความสอดคล้องและตรงกับความต้องการของผู้ใช้มากยิ่งขึ้น งานวิจัยนี้ได้นำเสนอโดเมนออนโทโลยีและกฎการปรับเปลี่ยนคำสั่งที่สามารถใช้ในการแปลงคำสั่งการสืบค้นที่มีความหมายซึ่งแสดงอยู่ในรูปแบบของภาษาสปาเก้ิล ให้เป็นภาษาสืบค้นสำหรับฐานข้อมูลซึ่งได้แก่ภาษาเอสคิวแอล ได้ ซึ่งจะทำให้ผลลัพธ์ที่สืบค้นได้นี้จะเป็นที่ตรงกับความต้องการของผู้ใช้อย่างแท้จริง และจากการทดลองสืบค้นความสอดคล้องระหว่างโครงสร้างของออนโทโลยีได้ค่าอัตราการรู้จำ เท่ากับ 93.99% ซึ่งหมายถึงผลจากการดึงเอกสารที่มีความครบถ้วนและตรงประเด็น จากการนำเอาออนโทโลยีมาช่วยในการเข้าถึงฐานข้อมูลแบบสื่อความหมายมีค่าความเชื่อมั่นเท่ากับ 93.99% เนื่องจากผลการแปลงองค์ประกอบที่สอดคล้องของฐานข้อมูลไปเป็นองค์ประกอบที่สอดคล้องของออนโทโลยีมีความซับซ้อนมาก อาจทำให้การอธิบายโครงสร้างของฐานข้อมูลบางส่วนไม่สัมพันธ์กัน แต่อย่างไรก็ตามการที่นำเอาออนโทโลยีมาช่วยในการสืบค้นข้อมูลนั้นมีประสิทธิภาพค่อนข้างดี ทำให้ผลลัพธ์ตามวัตถุประสงค์ที่กำหนดไว้ และสามารถแก้ปัญหาการสืบค้นฐานข้อมูลด้วยคำสำคัญ

เบรนดันและคณะ (Brendan et al, 2009) ได้ทำการวิจัย A Complete Translation from SPARQL into Efficient SQL ซึ่งได้นำเสนอ “SQL Model Base Algorithm” ซึ่งนำไปสร้างตัวดำเนินการ SPARQL Algebra ด้วยวิธีการเขียนโครงสร้างคำสั่งสืบค้นแบบเอสคิวแอลธรรมดาที่ไม่ซับซ้อน แต่สืบค้นฐานข้อมูลได้อย่างมีประสิทธิภาพ เพื่อแก้ปัญหาการสืบค้นเอกสารอาร์ดีเอฟใน Reasoner ที่ช้าลง เมื่อต้องเก็บข้อมูลไว้เป็นจำนวนมาก ถึงแม้ว่าเอกสารอาร์ดีเอฟนั้นจะได้มาจากการแปลงโครงสร้างฐานข้อมูลเชิงสัมพันธ์โดยที่ข้อมูลจัดเก็บอยู่แหล่งเดิมแล้วก็ตาม โดยลักษณะการแปลง

จะนำเอาคุณสมบัติการสืบค้นของภาษาสปาเกิ้ล ที่มีอยู่เดิมคือการเข้าถึงข้อมูลที่มีการจัดเก็บเป็น Triple ซึ่งประกอบด้วย Subject, Predicate และ Object ปรับให้เหมือนกับการเข้าถึงฐานข้อมูลของเอสคิวแอล ที่มุ่งเน้นการสืบค้นโดยคำสั่ง SELECT, FROM, WHERE และ ORDER BY พร้อมด้วยสนับสนุนส่วนขายแบบ DISTINCT, LIMIT และ OFFSET อีกด้วย นอกจากนั้นยังได้นำคำสั่งที่สร้างไปสืบค้นกลุ่มข้อมูลที่เป็นเว็บไซต์รูปภาพในสถานการณ์จริง งานวิจัยนี้มีข้อดีคือสามารถสืบค้น Triple pattern ที่มีความซับซ้อนได้ดี และทำได้อย่างรวดเร็ว

หวังและคณะ (Wang et al, 2009) ได้ทำการวิจัย Query Heterogeneous Relational Database using SPARQL ซึ่งศึกษาเกี่ยวกับปัญหาการเขียนคำสั่งการสืบค้นคืนใหม่ในกระบวนการสืบค้นแหล่งข้อมูลที่หลากหลายนอกจากนั้นยังได้วิเคราะห์ถึงความแตกต่าง และคล้ายคลึงกันระหว่างอาร์ดีเอฟ และ ER Model พร้อมทั้งอธิบายวิธีการแปลงจากโครงสร้างฐานข้อมูลไปเป็นอาร์ดีเอฟออนโทโลยี ที่นอกจากจะแปลงตารางเป็นคลาสแล้วก็ยังแปลงคอลัมน์ Primary Key ของตารางไปเป็นคลาสอีกด้วย ซึ่งในงานวิจัยนี้ยังใช้รูปแบบการอธิบายแหล่งข้อมูลบนพื้นฐานของสปาเกิ้ล ดังนี้

```
Table database1.proceeding can be described as the tuple
(       SELECT ?title, ?year
        WHERE  {?x pub_title ?title.
        ?X pub_year ?year}.
        {?title---> Title, ?year--->Year}
).
```

```
Table database1.rel can be described as the tuple
(       SELECT ?person, ?proceeding
        WHERE  {?x person Write ?proceeding}",
       {?person---> PersonId,
        ?proceeding---> ProceedingId}
).
```

จากรูปแบบการอธิบายแหล่งข้อมูลด้วยสปาเก้ล ข้างต้นได้นำมาปรับใช้เป็นหลักการเขียนคำสั่งการสืบค้นคืนใหม่ หรือ Query Rewriting Algorithm ที่กำหนดจากสองส่วนคือส่วนแรกเป็นการหาความสัมพันธ์ภายในตารางทั้งหมดแล้วนำมาแยกส่วนเป็น Minimal Connectable Unit (MCU) ซึ่งหมายถึงกลุ่มย่อยของ RDF Triple ตามข้อกำหนดของแหล่งจัดเก็บอธิบายไว้ ส่วนที่สองทำการเชื่อมโยง (JION) MCUs เพื่อสืบค้นเชิงความหมาตามรูปแบบของ RDF Graph Model ต่อไป โดยที่ Q เป็นคำสั่งสืบค้นสปาเก้ล, D_p เป็นการอธิบายแหล่งจัดเก็บตาราง D_p =(Q_p , μ) ซึ่ง P คือตาราง , μ คือความสัมพันธ์ และคำสั่งสืบค้นด้วยสปาเก้ล มีดังต่อไปนี้

```
SELECT ?title
WHERE {?X per_name “J.Wang”.
            ?X Write ?Y.
            ?Y title ?title. }
```

จากคำสั่งสืบค้นที่เขียนด้วยสปาเก้ล นำมาเขียนเป็นคำสั่งเอสคิวแอล ได้ดังนี้

```
SELECT proceeding.Title
FROM person,proceeding,rel
WHERE person.Name=’J. Wang’
AND person.PersonId=rel.PersionId
AND proceeding.ProceedingId=rel.ProceedingId
```

ทั้งคำสั่งสปาเก้ลและเอสคิวแอลต้องการสืบค้นชื่อบทความที่ตีพิมพ์โดยผู้วิจัยชื่อ “J.Wang”

จากเอกสารและงานวิจัยที่เกี่ยวข้อง จะพบว่ามีการพัฒนาและสร้างออนโทโลยีสำหรับการรวมข้อมูลเชิงความหมายของความรู้ในด้านต่าง ๆ เช่น ข้อมูลทางชีววิทยา ความรู้ด้านสมุนไพรไทย เป็นต้น และมีการวิจัยพัฒนาระบบสืบค้นข้อมูลเฉพาะเจาะจงโดยใช้คำเฉพาะและออนโทโลยี โดยปัญหาจะเกิดจากคำที่เขียนต่างกันแต่มีความหมายเหมือนกัน นอกจากนั้นยังพบปัญหาจากการศึกษางานวิจัยที่เกี่ยวข้องกับการจำแนก การจัดทำดัชนี และการกำหนดความสัมพันธ์ระหว่างข้อมูล ผู้วิจัยจึงเห็นความสำคัญของการวิจัยการเข้าถึงความรู้ทางด้านการเกษตรโดยใช้ออนโทโลยี ซึ่งงานวิจัยนี้จะมุ่งเน้นแก้ไขปัญหาดังกล่าวโดยการประยุกต์การค้นคืนเชิงความหมาย (Semantic Search) โดยนำเทคโนโลยี

เว็บเชิงความหมายมาจัดการความรู้ทางการเกษตร ซึ่งเจาะจงไปที่ไม้ผลเศรษฐกิจของประเทศไทย ได้แก่ลำไย ทุเรียน มังคุด และสามารถสืบค้นความรู้เชิงความหมาย

2.10 กรอบแนวคิดในการวิจัย

กรอบแนวคิดการวิจัยสรุปตามงานวิจัยที่เกี่ยวข้อง แสดงตามตารางที่ 2.10

ตารางที่ 2.10 กรอบแนวคิดในการวิจัย

ผู้วิจัย	ฐานความรู้			การค้นคืน		
	ขอบเขต	ออนโทโลยี	คลังคำ/ดัชนีคำสำคัญ	การวัดความคล้ายคลึงของคำค้น	วิธีการค้นคืน	ลำดับการนำเสนอผลลัพธ์
นาถธิชา เกษรพันธ์ (2554)	ผู้เชี่ยวชาญงานวิจัย	✓			คำสำคัญ	
เพ็ญพรรณ อัศวนพเกียรติ (2547)	ชีววิทยา	✓			ปัญญาประดิษฐ์	
กานดา ศรอินทร์ (2548)		✓	ดัชนีคำสำคัญ		บูลีน	แบบจำลองเวกเตอร์สเปซโมเดล
จารุณี แซ่หลี (2549)					คำสำคัญ	
นคร โคตรโสภา (2549)			ดัชนีคำสำคัญ		คำสำคัญที่มีการจัดลำดับ	การเปรียบเทียบความคล้ายคลึง
ชมพูนุช ภักติยานุวรรตน์ (2550)	โดเมนมหาวิทยาลัย	✓	การอธิบายข้อมูล		คำสำคัญ	
ประสิทธิ์ ประทุมรัตน์ (2550)	รูปภาพจดหมายเหตุ	✓	การอธิบายข้อมูล		ค้นคืนรูปภาพ	

ตารางที่ 2.10 กรอบแนวคิดในการวิจัย

ผู้วิจัย	ฐานความรู้			การค้นคืน		
	ขอบเขต	ออนโทโลยี	คลังคำ/ดัชนีคำสำคัญ	การวัดความคล้ายคลึงของคำค้น	วิธีการค้นคืน	ลำดับการนำเสนอผลลัพธ์
ภรศิษฐ์ กุหลาบทิพย์ (2550)			Wordnet/ดัชนีคำสำคัญ		คำสำคัญ	กำหนดน้ำหนักคำสำคัญ
วีรนันท์ มาลาศิลป์ (2550)	ข้อมูลบุคลากรและวิจัย	✓	การอธิบายข้อมูล		คำอธิบายข้อมูล	
บุญรัตน์ เผดิมรอด (2551)		✓			กลุ่มคำสำคัญ	
ปฏิคม ทองจริง (2552)	สมุนไพรไทย	✓			คำสำคัญ	
มนัสนันท์ ปัญญามี (2552)	เอกสารบนเว็บ		Wordnet		คำสำคัญและเว็บเพจ	
สมมณี ลูชะวงษ์ (2553)		✓			คำสำคัญ	
Brendan (2009)						
Wang (2009)						
งานวิจัยนี้	ไม้ผลเศรษฐกิจไทย	✓	คลังคำ	✓	กลุ่มคำสำคัญ	เปรียบเทียบความคล้ายคลึง

จากตารางที่ 2.10 ในงานวิจัยนี้มีกรอบแนวคิดการวิจัยสำหรับการค้นคืนความรู้ทางการเกษตร ที่แตกต่างจากงานวิจัยอื่นที่ได้ศึกษามา ดังนี้

- **การจัดการฐานความรู้**

การจัดการฐานความรู้ เป็นการจัดรูปแบบที่เหมาะสมสำหรับจัดเก็บความรู้ให้เป็นระบบ รวมถึงการจัดทำข้อมูลเพื่อช่วยในการค้นหาความรู้ ประเด็นสำคัญของการจัดการฐานความรู้ มีดังนี้

- *ขอบเขต*

งานวิจัยในปัจจุบันพบว่าจะมีความหลากหลายทางด้านโดเมนหรือขอบเขตของการจัดเก็บฐานความรู้ เนื่องจากแต่ละโดเมนหรือแต่ละขอบเขตจะมีการจัดการความรู้ที่แตกต่างกัน โดยโดเมนที่พบในงานวิจัยจะเป็นเรื่องเกี่ยวกับข้อมูลผู้เชี่ยวชาญ ข้อมูลมหาวิทยาลัย สมุนไพรไทย หรือเป็นโดเมนเฉพาะดังเช่นทางด้านชีววิทยา สำหรับงานวิจัยนี้มุ่งเน้นการจัดการความรู้ทางด้านเกษตรในเรื่องที่เกี่ยวกับไม้ผลเศรษฐกิจไทย

- *รูปแบบฐานความรู้*

สำหรับรูปแบบฐานความรู้โดยทั่วไปจะพบว่ามีการใช้รูปแบบออนโทโลยีในการจัดเก็บข้อมูลที่เป็นการอธิบายความรู้ เพื่อที่จะได้ใช้คำอธิบายข้อมูลดังกล่าวสำหรับการค้นคืนความรู้ ซึ่งออนโทโลยีที่พัฒนาขึ้นจะมีทั้งที่เป็นมาตรฐานสากล หรือพัฒนาเป็นการเฉพาะในงานวิจัยนั้น ๆ สำหรับงานวิจัยนี้จะใช้แนวทางของออนโทโลยีในการจัดเก็บคำอธิบายความรู้ โดยออนโทโลยีดังกล่าวจะพัฒนาตามแนวทางที่เป็นมาตรฐานสากล ได้แก่ AGROVOC Thesaurus และอรรถาภิธานเกษตรไทย

- *คลังคำ/ดัชนีคำสำคัญ*

การจัดทำคลังคำหรือดัชนีคำสำคัญก็เพื่อที่จะใช้สำหรับช่วยในการค้นคืนความรู้ในกรณีที่ผู้ใช้ระบุคำหลัก ระบบจะนำคำหลักดังกล่าวไปค้นหาว่ามีในคลังคำที่จัดทำหรือไม่ หากพบก็จะสามารถเชื่อมโยงไปถึงความรู้หรือเอกสารต่าง ๆ ได้ ซึ่งงานวิจัยนี้ได้จัดทำคลังคำโดยรวบรวมคำต่าง ๆ จากที่มีการจัดทำไว้เป็นมาตรฐาน และได้เพิ่มเติมคำที่เกี่ยวข้องจากเอกสารความรู้ด้านไม้ผลเศรษฐกิจไทยเพื่อที่จะช่วยให้การค้นคืนความรู้ดังกล่าวได้ตรงตามความต้องการของผู้ใช้มากขึ้น

- **การค้นคืนความรู้**

- *การวัดความคล้ายคลึงของคำค้น*

การค้นคืนความรู้โดยการใช้คำหลัก จะพบปัญหาของคำที่ผู้ใช้ระบุไม่ตรงกับคำหลักที่จัดเก็บไว้ในคลังคำ ซึ่งอาจมาจากการพิมพ์ผิด พิมพ์เกิน หรือพิมพ์ตกหล่น การวัดความคล้ายคลึงของคำค้นจึงเป็นการหาว่าคำค้นที่ผู้ใช้ระบุเหมือนกับคำหลักใดในคลังความรู้บ้าง เพื่อที่จะได้ใช้คำหลักดังกล่าวในการค้นหาความรู้หรือเอกสารที่เกี่ยวข้องต่อไป ในงานวิจัยนี้จะใช้สัมประสิทธิ์แจ็คการ์ด (Jaccard Index) ในการวัดความคล้ายคลึงของคำค้น

- *วิธีการค้นคืน*

การค้นหาความรู้หรือเอกสารที่เกี่ยวข้องหาใช้คำศัพท์ที่ได้จากผู้ใช้เพียงอย่างเดียวอาจจะทำให้การค้นค้นได้เอกสารความรู้ไม่ครบถ้วน เพราะถ้าหากเอกสารที่เกี่ยวข้องแต่ไม่ได้มีคำหลักตามที่ผู้ใช้กำหนด จะทำให้ไม่สามารถค้นคืนเอกสารความรู้นั้นได้ ดังนั้น วิธีการค้นคืนที่มีการกำหนดกลุ่มคำค้น

จากคำค้นที่ผู้ใช้ให้มา เพื่อใช้กลุ่มคำเหล่านั้นค้นหาเอกสารความที่เกี่ยวข้อง นอกจากนี้ ยังมีการประยุกต์การหาเส้นทางที่สั้นที่สุดสำหรับขั้นตอนวิธีการค้นหาเชิงความหมายในงานวิจัยนี้ด้วย

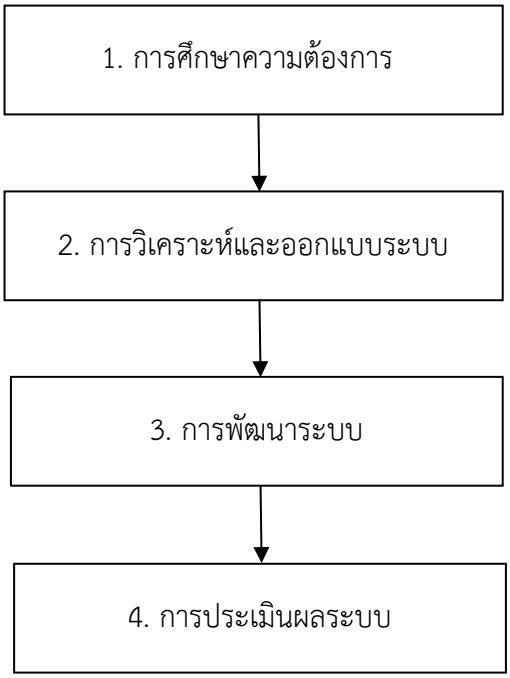
- *ลำดับการนำเสนอผลลัพธ์*

หลังจากที่การค้นหาได้ชุดของเอกสารที่เกี่ยวข้องมาแล้ว แต่เนื่องจากเอกสารมีจำนวนมาก จึงจำเป็นต้องมีการเรียงลำดับผลลัพธ์ โดยส่วนใหญ่แล้วจะเป็นการเรียงลำดับเพื่อให้สอดคล้องกับความต้องการของผู้ใช้ให้มากที่สุด ซึ่งอาจมีการใช้วิธีการที่แตกต่างกัน สำหรับงานวิจัยนี้จะมีการเสนอลำดับการนำเสนอผลลัพธ์ตามผลการเปรียบเทียบความคล้ายของเอกสารกับคำค้นจากผู้ใช้

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 3 วิธีการดำเนินการวิจัย

การดำเนินการวิจัยเป็นการวิจัยเชิงพัฒนา โดยมีขั้นตอนและวิธีการพัฒนาระบบประยุกต์ตามกระบวนการของวงจรการพัฒนาระบบ SDLC (System Development Life Cycle) โดยมีลำดับขั้นตอนการดำเนินการวิจัยดังดังภาพที่ 3.1



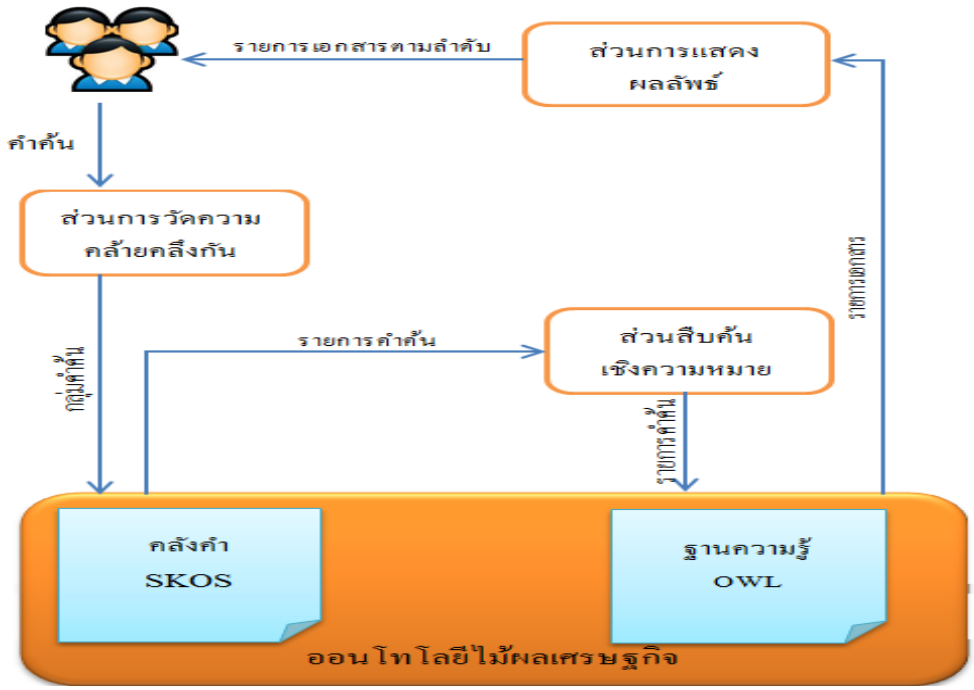
ภาพที่ 3.1 แสดงขั้นตอนการดำเนินการวิจัย

3.1 การศึกษาความต้องการ

การศึกษาข้อมูลเบื้องต้นจากเอกสารที่เกี่ยวข้อง จะประกอบไปด้วยข้อมูล 3 ส่วนที่สำคัญ ได้แก่ ข้อมูลการจัดเก็บและค้นคืนความรู้ ข้อมูลไม้ผลเศรษฐกิจของไทย และข้อมูลการวิจัยที่เกี่ยวข้อง เพื่อนำข้อมูลดังกล่าวมาวิเคราะห์และใช้เป็นแนวทางในการออกแบบระบบ ซึ่งผู้วิจัยได้มีการการศึกษาถึงปัญหาและรวบรวมข้อมูลในประเด็นของการสืบค้นความรู้และพฤติกรรมการสืบค้น ความรู้ด้านการเกษตรของไม้ผลเศรษฐกิจของประเทศไทย โดยมีการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องจากแหล่งสืบค้นข้อมูลงานวิจัยต่าง ๆ

3.2 การวิเคราะห์และออกแบบระบบ

จากการศึกษาและวิเคราะห์งานวิจัยและทฤษฎีที่เกี่ยวข้อง สามารถสรุปเป็นกรอบการทำงานของระบบการสืบค้นความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย ดังภาพที่ 3.2



ภาพที่ 3.2 ระบบการสืบค้นความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย

จากภาพที่ 3.2 กรอบการทำงานของระบบเริ่มจากผู้ใช้ต้องการสืบค้นข้อมูลจากอินเตอร์เน็ต ผู้ใช้จะป้อนคำสืบค้นผ่านทางเบราว์เซอร์ โดยติดต่อกับส่วนการวัดความคล้ายคลึงของคำค้นเพื่อทำการวัดความคล้ายคลึงกันของคำค้นกับคำหลักในคลังคำว่ามีความถูกต้องคำตามหลักไวยากรณ์หรือไม่ หลังจากนั้นจะได้รายการคำค้นที่ตรงหรือมีความคล้ายคลึงกันกับคำค้นส่งไปยังส่วนสืบค้นเชิงความหมายเพื่อทำการค้นหาเอกสารจากฐานความรู้ที่ตรงตามความต้องการของผู้ใช้ และมีส่วนแสดงผลลัพธ์ซึ่งแสดงรายการเอกสารที่มีการจัดลำดับความสำคัญไปยังผู้ใช้ โดยขั้นตอนวิธีการสืบค้นเชิงความหมายสำหรับเอกสารความรู้ด้านการเกษตร ประกอบไปด้วย 4 ส่วน ได้แก่ 1) ออนโทโลยีไม้ผลเศรษฐกิจ 2) ส่วนการวัดความคล้ายคลึงกันของคำค้น 3) ส่วนการสืบค้นเชิงความหมาย และ 4) ส่วนการแสดงผลลัพธ์

- **ออนโทโลยีไม้ผลเศรษฐกิจ**

การออกแบบออนโทโลยีไม้ผลเศรษฐกิจประกอบด้วยสองส่วน ได้แก่ ฐานความรู้ และคลังคำ ดังมีรายละเอียด ดังนี้

1) ฐานความรู้ เป็นการอธิบายรูปแบบโครงสร้างความสัมพันธ์ระหว่างข้อมูลโดยสร้างฐานความรู้ด้วยโปรแกรม Protégé โดยมีผลลัพธ์โครงสร้างฐานความรู้ออนโทโลยี ภายใต้ขอบเขตของไม้ผลเศรษฐกิจ ได้แก่ ทุเรียน มังคุด และลำไย โดยใช้โครงร่างพื้นฐานความสัมพันธ์ของเทอม (Term) จากอรรถาภิธานศัพท์เกษตรไทย เป็นอรรถาภิธาน (Thesaurus) ด้านการเกษตรในรูปอิเล็กทรอนิกส์ โดยคำศัพท์แต่ละคำมีการจัดความสัมพันธ์ในลักษณะลำดับชั้น (Hierarchical Relation) ให้มีเนื้อหาครอบคลุมข้อมูล ดังตารางที่ 3.1

ตารางที่ 3.1 แสดงรายการโครงสร้างข้อมูลของฐานความรู้ไม้ผลเศรษฐกิจ

1. ไม้ผลเศรษฐกิจ 1.1 ชื่อภาษาไทย 1.2 ชื่อภาษาอังกฤษ 1.3 วงศ์ 1.4 ลักษณะทั่วไป 1.5 แหล่งที่มา 1.6 ภาพ	**2. การปลูกและบำรุงรักษา** 2.1 วิธีปลูก 2.2 พื้นที่ปลูก 2.3 วิธีดูแลบำรุงรักษา 2.4 สภาพภูมิอากาศ	**3. โรคและแมลง** 3.1 ชื่อโรคหรือแมลง 3.2 อาการของโรคหรือแมลง 3.3 วิธีการป้องกันรักษาโรคหรือแมลง 3.4 ภาพของโรคหรือแมลง 3.5 แหล่งอ้างอิงของโรคหรือแมลง
4. สารเคมี 4.1 ชื่อสารออกฤทธิ์ 4.2 ชื่อสารสารออกฤทธิ์ทางการค้า 4.3 รายละเอียดการใช้สารเคมีทางการค้า 4.4 การแก้พิษสารเคมี 4.5 วิธีการเก็บสารเคมี	**5. การแปรรูป** 5.1 ชื่อผลิตภัณฑ์ 5.2 วิธีการแปรรูป 5.3 กระบวนการแปรรูป 5.4 ส่วนประกอบที่นำมาแปรรูป 5.5 รูปภาพผลิตภัณฑ์แปรรูป	**6. เครื่องจักรหรือเครื่องมือ** 6.1 ชื่อเครื่องจักร หรือเครื่องมือ 6.2 คุณสมบัติ 6.3 วิธีการใช้ 6.4 ภาพเครื่องจักร หรือเครื่องมือ 6.5 แหล่งที่มาเครื่องจักรหรือเครื่องมือ
7. การตลาดและการส่งออก 7.1 ชื่อตลาดซื้อขาย 7.2 ราคา 7.3 วันที่ 7.4 รายละเอียด	**8. เอกสารความรู้ไม้ผลเศรษฐกิจ**	

จากรายการโครงสร้างข้อมูลไม้ผลเศรษฐกิจ ได้นำมากำหนดคุณลักษณะและความสัมพันธ์ตามวิธีการออกแบบออนโทโลยี จากนั้นจะถูกแปลงให้อยู่ในรูปแบบของภาษาโอดัลเบิ้ลยู (Web Ontoloty Language) โดยใช้โปรแกรม Protégé

2) คลังคำ เป็นแหล่งรวบรวมคำต่าง ๆ โดยคำศัพท์แต่ละคำมีการจัดความสัมพันธ์ในลักษณะลำดับชั้น (Hierarchical Relation) ที่มีความหมายเหมือนกัน แต่เขียนต่างกัน โดยเน้นไปที่ คำหลัก คำพ้อง คำกว้างกว่า (Broader Term-BT) คำแคบกว่า (Narrower Term-NT) และ คำเกี่ยวข้อง (Related Term-RT) ซึ่งผู้วิจัยได้ประยุกต์ใช้ โดยใช้ตัวแบบเอสเคโอเอส (Simple Knowledge Organization System: SKOS) ซึ่งเป็นตัวแบบสำหรับสร้างคลังคำหรือมาตรฐานข้อมูลที่ใช้ร่วมกัน พื้นฐานของเอสเคโอเอส (SKOS) ประกอบไปด้วยความสัมพันธ์ ดังนี้ คือ คำหลัก (skos:preLabel) คำเหมือน (skos:altLabel) คำกว้างกว่า (skos:broader) คำแคบกว่า (skos:narrower) คำเกี่ยวข้อง (skos:related) (Fulvio and Laura, 2005)

- **ส่วนการวัดความคล้ายคลึงของคำค้น**

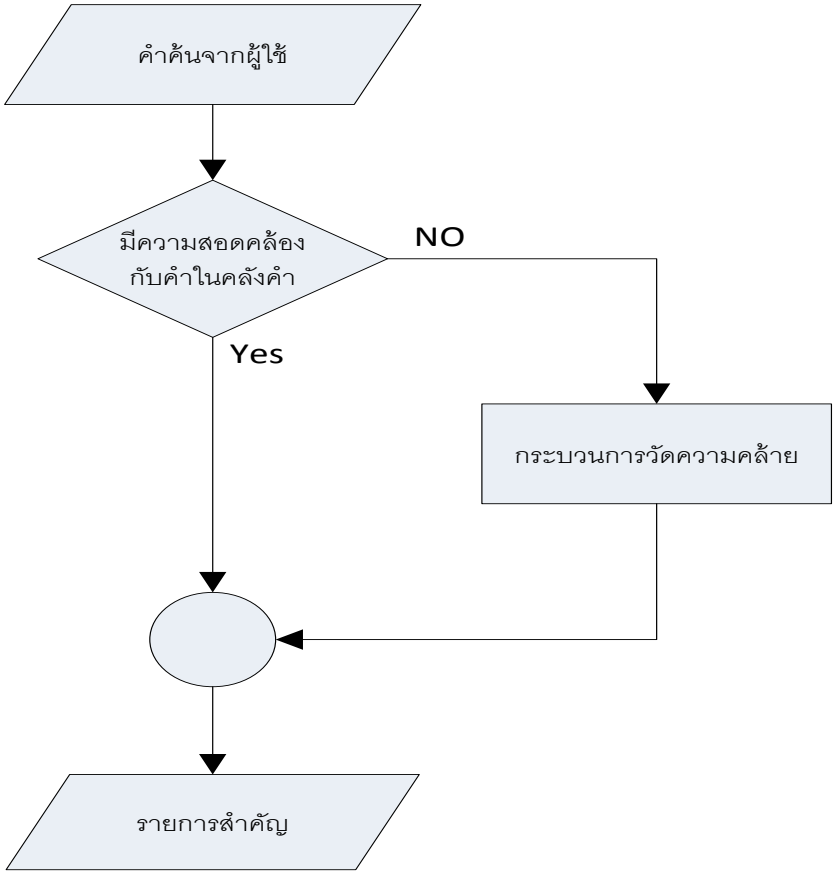
ในส่วนการวัดความคล้ายคลึงกันมีการใช้สัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ด (Jaccard Similarity Coeficient) สำหรับเปรียบเทียบความคล้ายคลึงคุณสมบัติของเซตข้อมูล การวัดความคล้ายคลึงของแจ็คการ์ดระหว่างเซตตัวอย่าง คือการหาผลหารของจำนวนของคุณสมบัติที่มีร่วมกันหารด้วยจำนวนของสมบัติที่มีทั้งหมด การวัดความคล้ายคลึงกันของคำค้น โดยผู้ใช้จะป้อนคำค้นเกี่ยวกับผลไม้เศรษฐกิจที่ต้องการเอกสารความรู้ จากนั้นระบบจะทำการสกัดหรือวัดความคล้ายคลึงกันของคำค้นว่าถูกต้องหรือไม่ในกรณีที่พิมพ์ถูก พิมพ์ผิด พิมพ์ตก และ พิมพ์เกิน เช่น คำว่า “ทุเรียน” ผู้ใช้อาจจะพิมพ์ผิดเป็น “ดุเทรียน” ระบบก็จะทำการคำนวณวัดความคล้ายคลึงกันระหว่างคำค้นกับคำที่เก็บอยู่ในคลังคำโดยใช้สัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ด แต่ในกรณีคำค้นที่พิมพ์ถูกต้องระบบจะทำการวัดความคล้ายคลึงระหว่างคำค้นกับคำที่เก็บอยู่ในคลังคำได้เลยว่าตรงกันกับคำค้นหรือไม่ ดังภาพที่ 3.3

กระบวนการวัดความคล้ายใช้สัมประสิทธิ์ความคล้ายของแจ๊คการ์ดในการวัดความคล้ายของคำค้นกับคำในเอกสารเพื่อระบุคำสำคัญ โดยสัมประสิทธิ์ความคล้ายของแจ๊คการ์ดให้ผลลัพธ์ในช่วง 0 ถึง 1 โดยที่ค่าความคล้ายยิ่งเข้าใกล้ 1 ยิ่งมีความคล้ายสูงมาก ดังสมการ 3.1

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \qquad (3.1)$$

A แทนเซตของตัวอักษรที่ไม่ซ้ำกันในคำที่ 1

B แทนเซตของตัวอักษรที่ไม่ซ้ำกันในคำที่ 2



ภาพที่ 3.3 ผังงานของการวัดความคล้ายคลึงของคำค้น

ในการวิจัยนี้ ใช้กระบวนการวัดความคล้ายใช้สัมประสิทธิ์ความคล้ายของแจ๊คการ์ดในการวัดความคล้ายของคำค้นกับคำในเอกสารเพื่อระบุคำสำคัญ โดยมีกระบวนการดำเนินการดังนี้

- *การเตรียมข้อมูลความสัมพันธ์ระหว่างคำ*

ในการวิจัยนี้มีการเตรียมข้อมูล 2 ส่วน คือ 1) การเตรียมข้อมูลที่เป็นคำ (ในที่นี้จะใช้คำว่า “คำ” แทนการหมายถึง กลุ่มคำหรือวลีด้วย) ซึ่งเป็นคำที่ถูกต้องตามหลักไวยากรณ์ โดยนำคำหลักมาจากอภิธานศัพท์เกษตรไทยที่เกี่ยวกับไม้ผลจำนวน 100 คำ 2) เตรียมข้อมูลเป็นคำที่ไม่ถูกต้องตามหลักไวยากรณ์ ซึ่งทดสอบใน 3 กลุ่ม ๆ ละ 30 คำ ได้แก่ คำสะกดผิด คำพิมพ์ขาด และ คำพิมพ์เกินกำหนดโดยผู้วิจัย แสดงตัวอย่างของคำปรากฏดังตารางที่ 3.2

ตารางที่ 3.2 แสดงตัวอย่างคำที่ปรากฏในดัชนี และคำสืบค้น (คำถูกตามไวยากรณ์ คำสะกดผิด คำพิมพ์ตก และคำพิมพ์เกิน)

คำสืบค้น	คำที่ปรากฏในดัชนี
คำถูกตามไวยากรณ์	ไม้ผลเศรษฐกิจ
ไม้ผลเศรษฐกิจ ทุเรียน มังคุด ลำไย ส้มโอ ส้มเขียวหวาน	ทุเรียน
	มังคุด
	ลำไย
คำสะกดผิด	ส้มโอ
ไม้ผบเศรษฐกิจ มุเรียน มัวคุด ลำใย ส้มโด ส้มเชียวหวาย	ส้มเขียวหวาน
	มะม่วง
	หมอนทอง
คำพิมพ์ตก	ก้านยาว
ไม้ผลเศรษฐกจ ทุเรีน มังคุ ลำย สมโอ ส้มเขียวหวน	เงาะ
	ลองกอง
	พันธุ์ชะนี
คำพิมพ์เกิน	พันธุ์ดอ
ไม้ผลเศรรษฐกิจ ทุเรียนน มังงคุด ลำไยบ ส้มโอ่ ส้มเขียววหวาน	พันธุ์พื้นเมือง
	กล้วย

- *การหาค่าความเกี่ยวข้องระหว่างคำสองคำด้วยสัมประสิทธิ์ของแจ็คการ์ด*

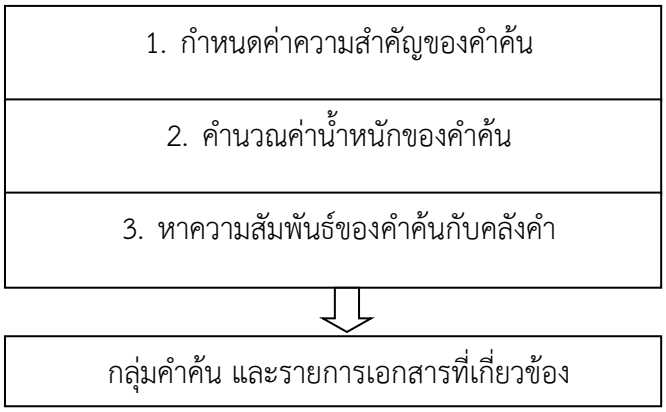
เพื่อการอธิบายให้ชัดเจนมากขึ้น ผู้วิจัยขอยกตัวอย่าง โดยกำหนดเซต A แทน “ส้มโอ” และเซต B แทน “ส้มเขียวหวาน” ซึ่งสามารถนำมาคำนวณค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดดังตัวอย่างการคำนวณสมการ 3.2

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|\{ส, ้, ม\}|}{|\{ส, ้, ม, โ, อ, เ, ข, ี, ย, ว, ห, ว, า, น\}|} = \frac{3}{13} = 0.23077 \qquad (3.2)$$

สัมประสิทธิ์ความคล้ายของแจ๊คการ์ดให้ผลลัพธ์ในช่วง 0 ถึง 1 ค่าความคล้ายที่เข้าใกล้ 1 จะมีความคล้ายสูงมาก ดังนั้นจากตัวอย่างข้างต้น สามารถสรุปได้ว่า เซต A และ เซต B มีค่าความคล้ายคลึง ที่ 0.23077 ซึ่งมีค่าความคล้ายที่ไม่สูง หรือตีความว่าไม่คล้าย โดยสัมประสิทธิ์ความคล้ายของแจ็คการ์ด (Jaccard’s Similiarity Coeeficient) สามารถนำมาใช้ในงานวิเคราะห์ความหมายของคำ โดยการอนุมานว่าคำเป็นเซตของตัวอักษรได้

- **ส่วนการสืบค้นเชิงความหมาย**

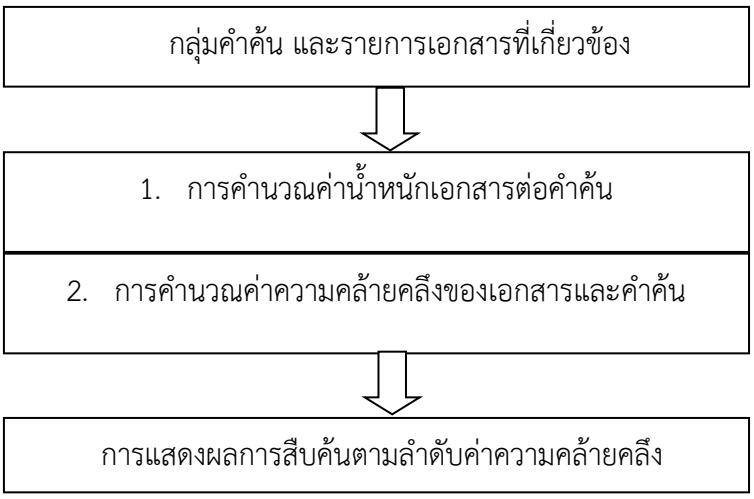
ส่วนการสืบค้นเชิงความหมาย เป็นส่วนสำคัญในการค้นหาความรู้หรือเอกสารจากฐานความรู้ไม้ผลเศรษฐกิจ โดยใช้คลังคำช่วยในการกำหนดคำหลักสำหรับการค้นหา โดยผลลัพธ์ที่ได้จะเป็นกลุ่มคำค้นและรายการเอกสารที่เกี่ยวข้อง ส่วนของการสืบค้นเชิงความหมาย ประกอบด้วยขั้นตอนที่สำคัญ 3 ขั้นตอน แสดงดังภาพที่ 3.4



ภาพที่ 3.4 แสดงขั้นตอนการสืบค้นเชิงความหมาย

- **ส่วนการแสดงผลลัพธ์**

ส่วนการแสดงผลลัพธ์เป็นการนำผลที่ได้จากขั้นตอนการสืบค้นเชิงความหมาย ซึ่งได้แก่กลุ่มคำค้นและรายการเอกสารที่เกี่ยวข้อง จากนั้นจะนำมาเปรียบเทียบความคล้ายคลึงของคำค้นและเอกสาร จากนั้นจะได้นำค่าที่ได้มาเรียงลำดับผลลัพธ์เพื่อนำเสนอต่อผู้ที่ต้องการค้นหาเอกสารความรู้ต่อไป สำหรับส่วนแสดงผลลัพธ์มีขั้นตอนการดำเนินการที่สำคัญ 2 ขั้นตอน แสดงดังภาพที่ 3.5



ภาพที่ 3.5 แสดงขั้นตอนส่วนแสดงผลลัพธ์การสืบค้น

3.3 การพัฒนาระบบ

ในส่วนของการพัฒนาระบบสืบค้นเชิงความหมาย ผู้วิจัยได้ทำการพัฒนาระบบโดยใช้เทคนิคการวัดความคล้ายคลึง ของแจ๊คการ์ด และขั้นตอนวิธีการสืบค้นเชิงความหมายด้วยภาษา PHP และใช้ RAP (RDF API for PHP) เพื่อเชื่อมต่อกับฐานความรู้เกษตรโดยใช้ตัวแบบออนโทโลยี (สำหรับการติดตั้งโปรแกรม RAP ดูรายละเอียดได้ที่ภาคผนวก ข) ใช้ภาษา สปาเก้ล Query Language สำหรับการค้นหาความรู้ ซึ่งผู้ใช้ต้องระบุคำสำคัญที่ต้องการค้นหา ระบบจะทำการวัดความคล้ายคลึงกันของคำค้นกับคำสำคัญที่เก็บอยู่ในคลังคำเป็นอันดับแรก และได้คำสำคัญสำหรับนำไปสืบค้นเชิงความหมาย เพื่อทำการค้นข้อมูลเอกสารในไฟล์โอดัลเบิ้ลยูแอล ที่เป็นฐานความรู้ออนโทโลยีตามขั้นตอนวิธีการสืบค้นเชิงความหมายสำหรับเอกสารความรู้ผลไม้เศรษฐกิจ โดยแสดงผลลัพธ์ผ่านเว็บเบราว์เซอร์ ที่ประกอบด้วย ชื่อเอกสาร ชื่อผู้เขียน รายละเอียดโดยย่อ และจัดลำดับแสดงเปอร์เซ็นต์ความถูกต้องของผลลัพธ์ในการสืบค้นเชิงความหมายให้ตรงกับความต้องการของผู้ใช้

3.4 การประเมินผลระบบ

สำหรับการประเมินระบบประกอบด้วยสองส่วน ได้แก่ การประเมินความถูกต้องของระบบ และการประเมินความสามารถในการทำงานของระบบ โดยมีรายละเอียดในการประเมินแต่ละส่วน ดังนี้

- **การประเมินความถูกต้องของระบบ**

สำหรับการประเมินความถูกต้องของระบบสืบค้นเชิงความหมายเปรียบเทียบกับการกำหนดความสำคัญของเอกสารต่อคำค้นของผู้เชี่ยวชาญ โดยใช้วิธีวัดค่ามาตรฐานในการประเมินความถูกต้อง 3 ค่า ได้แก่ ค่าความแม่นยำ (Precision) ค่าการจำได้ (Recall) และอัตราการรู้จำ (F-measure) โดยประยุกต์จากสมการ 2.1, 2.2 และ 2.3 ดังแสดงไว้ในหน้า 40-42 (Miao et al, 2009)

Precision คือค่าความแม่นยำ ที่จะแสดงให้เห็นว่าระบบที่พัฒนาขึ้นมามีความแม่นยำเพียงใด โดยสามารถคำนวณได้จากสูตรดังสมการ 3.3

$$Precision = \frac{TP}{TP+FP} \times 100\% \qquad (3.3)$$

TP (True Prositive) แทน เอกสารที่เลือก และถูกเลือกโดยผู้เชี่ยวชาญ

FP (False Prositive) แทน เอกสารที่เลือก แต่ไม่ถูกเลือกโดยผู้เชี่ยวชาญ

Recall คือค่าการจำได้ ที่แสดงให้เห็นว่าเมื่อระบบได้ทำการดึงคำตอบออกมาแล้วมีความถูกต้องเพียงใด โดยสามารถคำนวณได้จากสูตรดังสมการ 3.4

$$Recall = \frac{TP}{TP+FN} \times 100\% \qquad (3.4)$$

TP (True Prositive) แทน เอกสารที่เลือก และถูกเลือกโดยผู้เชี่ยวชาญ

FN (False Negative) แทน เอกสารที่ไม่เลือก แต่ถูกเลือกโดยผู้เชี่ยวชาญ

F-measure คืออัตราการรู้จำ ซึ่งเป็นค่าเฉลี่ยที่ให้ความสำคัญกับความแม่นยำและค่าการจำได้ คำนวณได้จากสูตรดังสมการ 3.5

$$F = (2 \text{ x Precision x Recall}) / (\text{Precision} + \text{Recall}) \quad (3.5)$$

- **การประเมินความสามารถของระบบ**

การประเมินความสามารถของระบบสืบค้นสำหรับการเข้าถึงความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย ผู้วิจัยได้ทำการสร้างแบบประเมินที่ใช้ในการวิจัยครั้งนี้ ดังนี้

1. สร้างแบบประเมินซึ่งเป็นแบบสอบถาม (Questionnaire) แบ่งออกเป็น 2 ตอน ดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ลักษณะแบบสอบถามเป็นแบบตรวจสอบรายการ (Checklist) จำนวน 2 ข้อ ประกอบด้วย เพศ และ ระดับการศึกษา

ตอนที่ 2 ความคิดเห็นเกี่ยวกับความสามารถของระบบ ลักษณะแบบสอบถามเป็นแบบมาตราส่วนประมาณค่า (Rating Scale) จำนวน 20 ข้อ โดยมีเกณฑ์การให้คะแนน ตามวิธีการของลิเคอร์ท (Likert) มี 5 ระดับ (บุญชม ศรีสะอาด. 2545 : 102-103) ดังนี้

5 หมายถึง เหมาะสมมากที่สุด

4 หมายถึง เหมาะสมมาก

3 หมายถึง เหมาะสมปานกลาง

2 หมายถึง เหมาะสมน้อย

1 หมายถึง เหมาะสมน้อยที่สุด

โดยครอบคลุมข้อมูลประเมินความสามารถของระบบในด้านต่าง ๆ ดังนี้

1) Functional requirement test เป็นการประเมินความสามารถของระบบว่าตรงตามความต้องการมากน้อยเพียงใด จำนวน 4 ข้อ

2) Functional test เป็นการประเมินความถูกต้องของระบบว่าสามารถทำงานได้ถูกต้องตรงตามหน้าที่มากน้อยเพียงใด จำนวน 4 ข้อ

3) Usability test เป็นการประเมินลักษณะการใช้งานของระบบว่ามีความง่ายต่อการใช้งานมากน้อยเพียงใด จำนวน 6 ข้อ

4) Performance test เป็นการประเมินประสิทธิภาพของระบบว่ามีประสิทธิภาพมากน้อยเพียงใด จำนวน 3 ข้อ

5) Security test เป็นการประเมินด้านการรักษาความปลอดภัยของข้อมูลว่ามีมากน้อยเพียงใด จำนวน 3 ข้อ

2. นำแบบประเมินที่สร้างขึ้นเสนอต่อผู้เชี่ยวชาญตรวจสอบความสอดคล้องระหว่างข้อคำถามกับจุดประสงค์ของการประเมิน โดยประเมินค่าดัชนีความสอดคล้อง (Item-Objective Congruence Index : IOC) (กรมวิชาการ. 2545 : 65) ซึ่งผู้เชี่ยวชาญประกอบด้วย

2.1 รองศาสตราจารย์ สายตา บุญโฉม คณบดีฝ่ายบริหาร คณะเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน (ผู้เชี่ยวชาญด้านงานวิจัย)

2.2 รองศาสตราจารย์ ชาติ วรรณกุล รองคณบดีฝ่ายบริหาร คณะเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน (ผู้เชี่ยวชาญด้านคอมพิวเตอร์)

2.3 ผู้ช่วยศาสตราจารย์ ดร. อำภาศรี พ่อค้า อาจารย์ประจำอาจารย์ คณะเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน (ผู้เชี่ยวชาญด้านงานวิจัยและสถิติ)

สำหรับการประเมินค่าดัชนีความสอดคล้องข้อคำถามกับจุดประสงค์การประเมิน (IOC) มีเกณฑ์การใช้คะแนน ดังนี้

ให้คะแนน +1 เมื่อแน่ใจว่าข้อคำถามนั้นประเมินตามจุดประสงค์

ให้คะแนน 0 เมื่อไม่แน่ใจว่าคำถามนั้นประเมินตามจุดประสงค์

ให้คะแนน -1 เมื่อแน่ใจว่าคำถามนั้นไม่สอดคล้องกันระหว่างคำถามกับจุดประสงค์

3. วิเคราะห์ข้อมูลการหาค่าดัชนี ความสอดคล้องระหว่างข้อคำถามของแบบประเมินกับจุดประสงค์ในการประเมินความสามารถของระบบ ผลการคำนวณค่า IOC อยู่ระหว่าง 0.67-1.00 (รายละเอียดปรากฏตามภาคผนวก ค) ซึ่งผ่านเกณฑ์ที่กำหนดให้ข้อคำถามที่มีค่าคะแนน IOC ตั้งแต่ 0.5 ขึ้นไปสามารถใช้เป็นข้อคำถามได้ ในการคำนวณค่า ICO ใช้สูตรดังสมการ 3.6 ดังนี้ (กรมวิชาการ, 2545 : 65)

$$IOC = \frac{\Sigma R}{N} \qquad (3.6)$$

IOC คือ ดัชนีความสอดคล้องระหว่างคำถามกับจุดประสงค์

R คือ คะแนนของผู้เชี่ยวชาญ

ΣR คือ ผลรวมของคะแนนผู้เชี่ยวชาญแต่ละคน

N คือ จำนวนผู้เชี่ยวชาญ

4. นำแบบประเมินที่ผ่านการประเมินค่า IOC แล้วไปให้ผู้เชี่ยวชาญ (ที่ไม่ใช่ข้อ 2.1-2.3) ทำการประเมินความสามารถในการทำงานของระบบสืบค้น จำนวน 15 คน ดังนี้ 1) ผู้เชี่ยวชาญด้านคอมพิวเตอร์จำนวน 5 คน 2) ผู้เชี่ยวชาญทางด้านการเกษตร จำนวน 10 คน ประกอบด้วยบุคลากรที่สอนสาขาเกี่ยวข้องกับการเกษตรด้านไม้ผลในสถาบันการศึกษา จำนวน 5 คน และนักวิชาการเกษตรระดับจังหวัด จำนวน 5 คน

5. ผู้วิจัยนำแบบประเมินความสามารถของระบบสืบค้น ที่ได้จากผู้เชี่ยวชาญจำนวน 15 คน มาวิเคราะห์ระดับความเหมาะสมโดยใช้สถิติค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน และสำหรับการวิเคราะห์จะใช้ค่าเฉลี่ยเทียบกับเกณฑ์การประเมินซึ่งกำหนดโดยผู้วิจัย ดังตารางที่ 3.3

ตารางที่ 3.3 เกณฑ์การประเมินการทำงานของระบบ

เกณฑ์การให้คะแนน (ค่าเฉลี่ย)	การแปลผล
4.50 – 5.00	ดีมาก
3.50 - 4.49	ดี
2.50 – 3.49	ปานกลาง
1.50 – 2.49	พอใช้
1.00 – 1.49	ต้องปรับปรุง

สำหรับตัวอย่างแบบสอบถามและผลการประเมินของผู้เชี่ยวชาญปรากฏตามภาคผนวก ค

บทที่ 4 ผลการวิจัย

การดำเนินการวิจัยครั้งนี้ ผู้วิจัยเสนอผลการวิจัยและพัฒนาระบบสืบค้นเชิงความหมายสำหรับเอกสารความรู้ด้านการเกษตรสำหรับไม้ผลเศรษฐกิจของประเทศไทย ดังนี้

4.1 ออนโทโลยีไม้ผลเศรษฐกิจ

4.2 การวัดความคล้ายคลึง

4.3 การค้นหาเชิงความหมาย

4.4 การแสดงผลลัพธ์

4.5 ผลการประเมินระบบ

4.6 การอภิปรายผล

4.1 ออนโทโลยีไม้ผลเศรษฐกิจ

ผลการออกแบบออนโทโลยีไม้ผลเศรษฐกิจ มีดังนี้

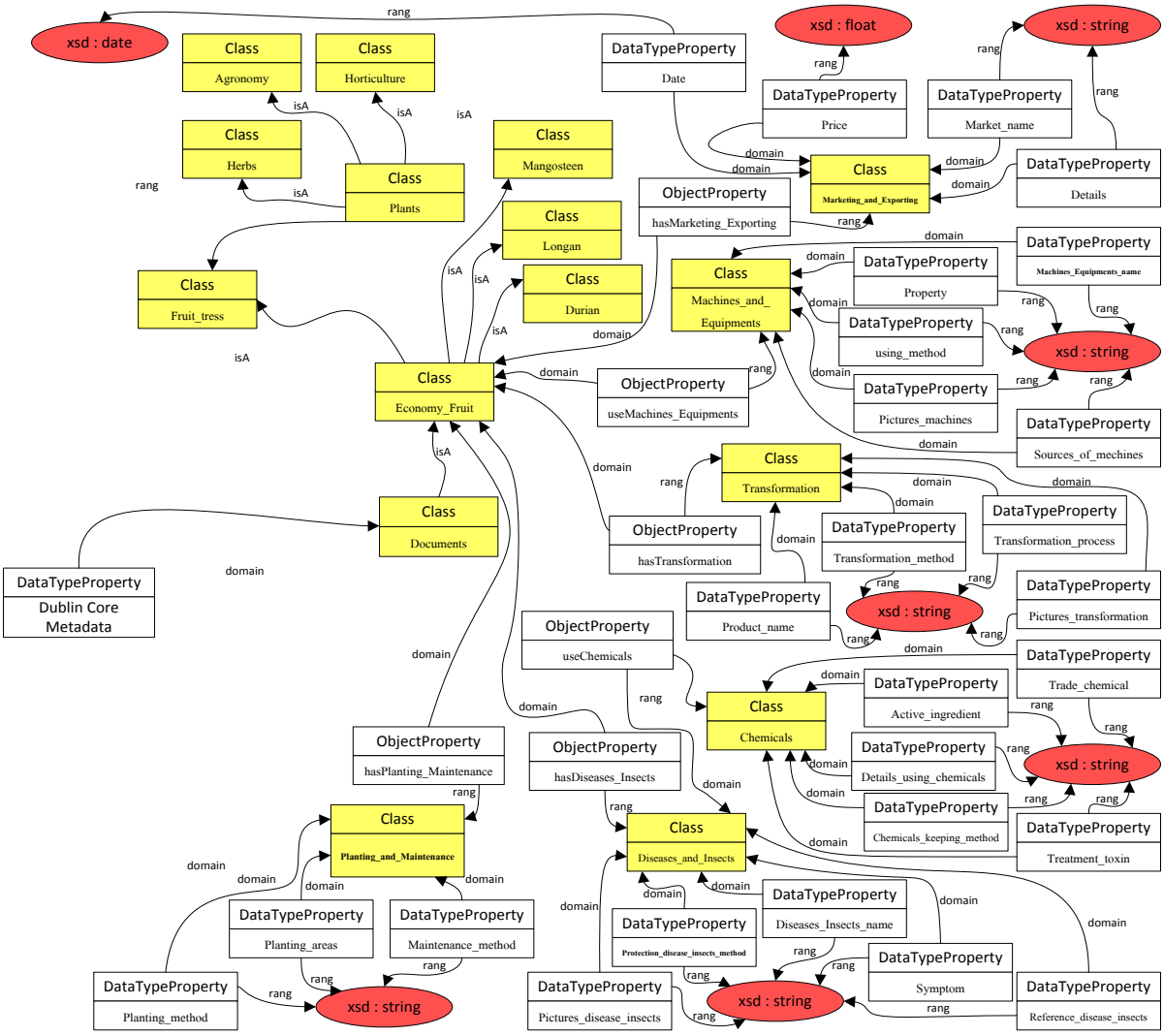
- **ฐานความรู้**

จากการออกแบบฐานความรู้ด้วยตัวแบบออนโทโลยี ได้ออกแบบตามโครงสร้างข้อมูลไม้ผลเศรษฐกิจตามตารางที่ 3.1 รายละเอียดของคลาสที่กำหนดขึ้นมาจำนวน 8 คลาส ปรากฏดังตารางที่ 4.1

ตารางที่ 4.1 รายละเอียดของคลาสในฐานความรู้ไม้ผลเศรษฐกิจ

คลาส	รายละเอียด
Economy_Fruit	คลาสแทนข้อมูลไม้ผลเศรษฐกิจ
Planting_and_Maintenance	คลาสแทนข้อมูลการปลูกและบำรุงรักษา
Diseases_and_Insects	คลาสแทนข้อมูลโรคและแมลง
Chemicals	คลาสแทนข้อมูลสารเคมี
Transformation	คลาสแทนข้อมูลการแปรรูป
Machines_and_Equipments	คลาสแทนข้อมูลเครื่องจักรหรือเครื่องมือ
Marketing_and_Exporting	คลาสแทนข้อมูลการตลาดและการส่งออก
Documents	คลาสแทนข้อมูลเอกสารความรู้ไม้ผลเศรษฐกิจ

หลังจากออกแบบคลาสแล้ว ได้กำหนดออฟเจคพร็อพเพอร์นี้และดาตาไทป์พร็อพเพอร์ตี้ พร้อมกับกำหนดความสัมพันธ์ระหว่างคลาสและพร็อพเพอรีตี้เหล่านั้น ผลการออกแบบฐานความรู้ไม้ผลเศรษฐกิจไทยโดยใช้ตัวแบบออนโทโลยีแสดงดังภาพที่ 4.1



ภาพที่ 4.1 โครงสร้างฐานความรู้ไม้ผลเศรษฐกิจโดยใช้ตัวแบบออนโทโลยี

จากภาพที่ 4.1 รายละเอียดออฟเจคพร็อพเพอร์ตี้ (Object Properties) ที่ได้กำหนดไว้จำนวน 9 ออฟเจคพร็อพเพอร์ตี้แสดงตามตารางที่ 4.2

ตารางที่ 4.2 รายละเอียดของออฟเจคพร๊อพเพอร์ตี้ในฐานความรู้ไม้ผลเศรษฐกิจ

พร๊อพเพอร์ตี้	รายละเอียด
hasDocument	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส คลังคำ (skos:Concept) และ Document
hasTerms	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Document และคลังคำ (skos:Concept)
hasDiseases_Insects	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Economy_Fruit และ Diseases_and_Insects
hasMarketing_Exporting	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Economy_Fruit และ Marketing_and_Exporting
hasPlanting_Maintenance	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Economy_Fruit และ Planting_and_Maintenance
hasTransformation	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Economy_Fruit และ Transformation
useChemicals	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Diseases_and_Insects และ Chemicals
useEquipments	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Transformation และ Machines_and_Equipments
useMachines_Equipments	เป็นพร๊อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส Planting_and_Maintenance และ Machines_and_Equipments

จากภาพที่ 4.1 รายละเอียดดาต้าไท้ป์พร๊อพเพอร์ตี้ (DataType Properties) จำนวน 34 ดาต้าไทป์พร็อพเพอร์ตี้ แสดงตามตารางที่ 4.3

ตารางที่ 4.3 รายละเอียดของดาต้าไท้ป์พร๊อพเพอร์ตี้ในฐานความรู้ไม้ผลเศรษฐกิจ

พร๊อพเพอร์ตี้	ประเภทข้อมูล	รายละเอียด
inThai	String	เป็นการระบุค่าคงที่แสดงชื่อภาษาไทยไม้ผลเศรษฐกิจ
inEnglish	String	เป็นการระบุค่าคงที่แสดงชื่อภาษาอังกฤษไม้ผลเศรษฐกิจ
Lineage	String	เป็นการระบุค่าคงที่แสดงชื่อวงศ์ไม้ผลเศรษฐกิจ
Generalization	String	เป็นการระบุค่าคงที่แสดงลักษณะทั่วไปไม้ผลเศรษฐกิจ

ตารางที่ 4.3 รายละเอียดของดาต้าไท้ป์พร๊อพเพอร์ตี้ในฐานความรู้ไม้ผลเศรษฐกิจ

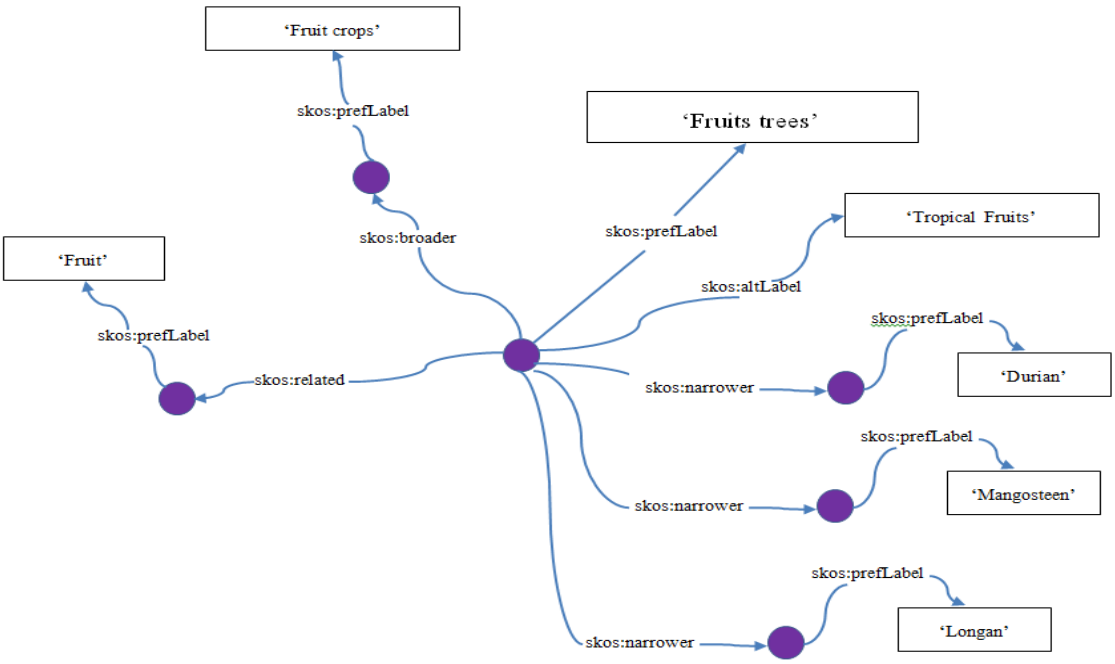
พร๊อพเพอร์ตี้	ประเภทข้อมูล	รายละเอียด
Sources	String	เป็นการระบุค่าคงที่แสดงแหล่งที่มาไม้ผลเศรษฐกิจ
Pictures	String	เป็นการระบุค่าคงที่แสดงภาพไม้ผลเศรษฐกิจ
Planting_method	String	เป็นการระบุค่าคงที่แสดงวิธีปลูก
Planting_areas	String	เป็นการระบุค่าคงที่แสดงพื้นที่ปลูก
Maintenance_method	String	เป็นการระบุค่าคงที่แสดงวิธีดูแลบำรุงรักษา
Weather	String	เป็นการระบุค่าคงที่แสดงสภาพภูมิอากาศ
Diseases_Insects_name	String	เป็นการระบุค่าคงที่แสดงชื่อโรคหรือแมลง
Symptom	String	เป็นการระบุค่าคงที่แสดงอาการของโรคหรือแมลง
Protection_disease_insects_method	String	เป็นการระบุค่าคงที่แสดงวิธีการป้องกันรักษาโรคหรือแมลง
Pictures_disease_insects	String	เป็นการระบุค่าคงที่แสดงภาพของโรคหรือแมลง
Reference_disease_insects	String	เป็นการระบุค่าคงที่แสดงแหล่งอ้างอิงของโรคหรือแมลง
Active_ingredient	String	เป็นการระบุค่าคงที่แสดงชื่อสารออกฤทธิ์
Trade_chemical	String	เป็นการระบุค่าคงที่แสดงชื่อสารสารออกฤทธิ์ทางการค้า
Details_using_chemicals	String	เป็นการระบุค่าคงที่แสดงรายละเอียดการใช้สารเคมีทางการค้า
Treatment_toxin	String	เป็นการระบุค่าคงที่แสดงการแก้พิษสารเคมี
Chemicals_keeping_method	String	เป็นการระบุค่าคงที่แสดงวิธีการเก็บสารเคมี
Product_name	String	เป็นการระบุค่าคงที่แสดงชื่อผลิตภัณฑ์การแปรรูป
Transformation_method	String	เป็นการระบุค่าคงที่แสดงวิธีการแปรรูป
Transformation_process	String	เป็นการระบุค่าคงที่แสดงกระบวนการแปรรูป
Part_of_transformation	String	เป็นการระบุค่าคงที่แสดงส่วนประกอบที่นำมาแปรรูป
Pictures_transformation	String	เป็นการระบุค่าคงที่แสดงรูปภาพผลิตภัณฑ์แปรรูป
Machines_Equipments_name	String	เป็นการระบุค่าคงที่แสดงชื่อเครื่องจักรหรือเครื่องมือ
Property	String	เป็นการระบุค่าคงที่แสดงคุณสมบัติเครื่องมือ

ตารางที่ 4.3 รายละเอียดของดาต้าไท้ป์พร๊อพเพอร์ตี้ในฐานความรู้ไม้ผลเศรษฐกิจ

พร๊อพเพอร์ตี้	ประเภทข้อมูล	รายละเอียด
using_method	String	เป็นการระบุค่าคงที่แสดงวิธีการใช้เครื่องมือ
Pictures_machines	String	เป็นการระบุค่าคงที่แสดงภาพเครื่องจักรหรือเครื่องมือ
Sources_of_mechines	String	เป็นการระบุค่าคงที่แสดงแหล่งที่มาเครื่องจักร หรือเครื่องมือ
Market_name	String	เป็นการระบุค่าคงที่แสดงชื่อตลาดซื้อขาย
Price	Float	เป็นการระบุค่าคงที่แสดงราคาในตลาด
Date	Date	เป็นการระบุค่าคงที่แสดงวันที่ของราคาในตลาด
Details	String	เป็นการระบุค่าคงที่แสดงรายละเอียดการตลาด

- **คลังคำ**

ตัวอย่างการนิยามศัพท์ผลไม้เศรษฐกิจใน SKOS ดังภาพที่ 4.2 โดยโครงสร้างความสัมพันธ์ของคำแสดงให้เห็นว่า ไม้ผลยืนต้น (Fruits trees) มีผลไม้ (Fruit) เป็นคำเหมือน (skos:related) มีไม้ผล (Fruit crops) เป็นคำที่กว้างกว่า (skos:broader) มีผลไม้เขตร้อน (Topical Fruits) เป็นคำที่เกี่ยวข้อง (skos:altLabel) และมีทุเรียน (Durian) มังคุด (Mangosteen) และ ลำไย (Longan) เป็นคำแคบกว่า (skos:narrower)



ภาพที่ 4.2 แสดงการนิยามศัพท์ผลไม้เศรษฐกิจในเอสเคโอเอส

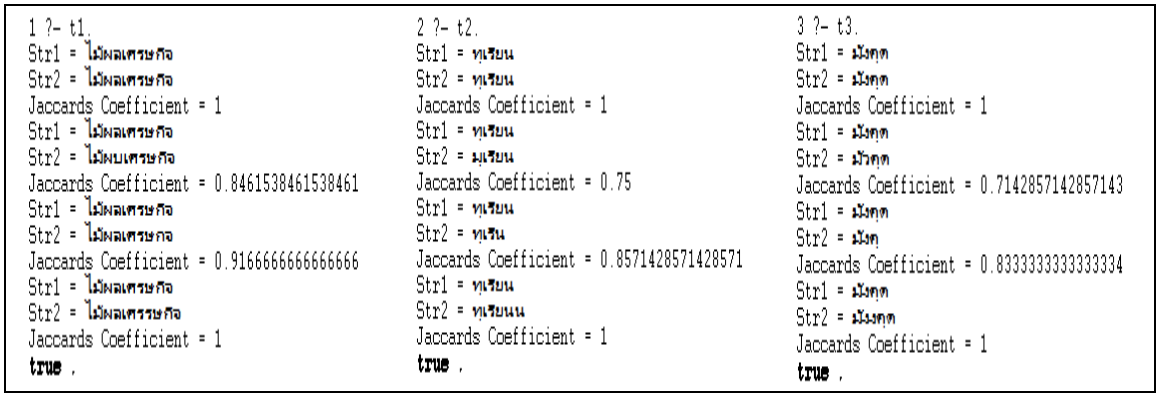
4.2 การวัดความคล้ายคลึงของคำค้น

กระบวนการวัดความคล้ายโดยใช้สัมประสิทธิ์ความคล้ายของแจ๊คการ์ดในการวัดความคล้ายของคำค้นกับคำในเอกสารเพื่อระบุคำสำคัญ (Niwattanakul et al, 2013) มีรายละเอียดในการดำเนินการดังขั้นตอนต่อไปนี้

1. **ทดสอบสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ด** ด้วยโปรแกรมสำหรับภาษาโปรล็อก ซึ่งภาษาโปรล็อกเป็นภาษาเชิงอนุมาน (Inference Engine) กล่าวคือเป็นโปรแกรมที่เรียนรู้ได้จากสิ่งที่นักพัฒนากำหนดให้ หรืออีกในหนึ่งคือเป็นภาษาแบบปัญญาประดิษฐ์ ซึ่งมีไวยากรณ์ที่สามารถเรียนรู้ได้ด้วยระยะเวลาอันสั้น เหมาะสำหรับการพัฒนาโปรแกรมแก้ปัญหาเชิงตรรกะ ปัญญาประดิษฐ์ และภาษาศาสตร์เชิงคำนวณ ซึ่งผลการวัดความคล้ายคลึงของคำค้น ปรากฏดังตัวอย่างโค้ดดังภาพที่ 4.3 และตัวอย่างการทำงานของโปรแกรมดังภาพที่ 4.4

```
jaccardsCoefficient(T1,T2,R) :- name(T1, CharList1),
                                name(T2, CharList2),
                                intersection(CharList1,CharList2,RL),
                                sizeList(RL,IR),
                                union(CharList1,CharList2,RL2),
                                sizeList(RL2,UR),
                                R is (IR)/(UR).

union([],L,R)    :-  union(L,[],R).
union([H|TL],L,R)    :-  clearDuplicate(H,TL,NTL),
                         clearDuplicate(H,L,NL),
                         union(NTL,NL,RL),
                         R = [H|RL].

intersection([],_,[]).
intersection([H|TL],CharList2,R)    :-  (   inList(H,CharList2)
                                        ->  clearDuplicate(H,CharList2,NewCharList2),
                                            intersection(TL,NewCharList2,TR),
                                            R = [H|TR]
                                        ;   NewCharList2 = CharList2,
                                            intersection(TL,NewCharList2,TR),
                                            R = TR  ).

clearDuplicate(_,[],[]).
clearDuplicate(C,[H|TL],R)  :-  clearDuplicate(C,TL,TR),
                                (   C = H
                                    ->  R = TR
                                    ;   R = [H|TR]  ).

sizeList([],0).
sizeList([H|TL],R)  :-  clearDuplicate(H,TL,TR),
                        sizeList(TR,RT),
                        R is RT + 1.
```

ภาพที่ 4.3 โค้ดโปรแกรมการคำนวณสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดในภาษาโปรล็อก



ภาพที่ 4.4 ตัวอย่างแสดงการประมวลผลโปรแกรม

2. การทดสอบเปรียบเทียบคำแบบปกติ เพื่อทดสอบค่าความถูกต้องของสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลกับคำที่ถูกต้องตามไวยากรณ์ แสดงตัวอย่างผลในตารางที่ 4.4

ตารางที่ 4.4 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลกับคำที่ถูกต้องตามไวยากรณ์

	ไม้ผลเศรษฐกิจ	ทุเรียน	มังคุด	ลำไย	ส้มโอ	ส้มเขียวหวาน	มะม่วง
ไม้ผลเศรษฐกิจ	1						
ทุเรียน	0.118	1					
มังคุด	0.059	0.083	1				
ลำไย	0.143	0.1	0	1			
ส้มโอ	0.133	0	0.1	0	1		
ส้มเขียวหวาน	0.15	0.286	0.063	0.071	0.231	1	
มะม่วง	0.063	0	0.222	0	0.111	0.143	1

3. การทดสอบเปรียบเทียบคำแบบมีข้อผิดพลาด เพื่อทดสอบค่าความถูกต้องของสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลกับคำที่สะกดผิด พิมพ์ตก และพิมพ์เกิน แสดงตัวอย่างการคำนวณดังตารางที่ 4.4 ถึงตารางที่ 4.9

ตารางที่ 4.5 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “ไม้ผลเศรษกิจ”

	พิมพ์ถูก ไม้ผลเศรษฐกิจ	พิมพ์ผิด ไม้ผบเศรษฐกิจ	พิมพ์ตก ไม้ผลเศรษฐกจ	พิมพ์เกิน ไม้ผลเศรรษฐกิจ
ไม้ผลเศรษฐกิจ	1	0.786	0.846	0.923
ทุเรียน	0.118	0.111	0.118	0.111
มังคุด	0.059	0.056	0.059	0.056
ลำไย	0.143	0.063	0.143	0.133
ส้มโอ	0.133	0.125	0.133	0.125
ส้มเขียวหวาน	0.15	0.143	0.15	0.149
มะม่วง	0.063	0.059	0.063	0.059

ตารางที่ 4.6 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “ทุเรียน”

	พิมพ์ถูก ทุเรียน	พิมพ์ผิด มุเรียน	พิมพ์ตก ทุเรีน	พิมพ์เกิน ทุเรียนน
ไม้ผลเศรษฐกิจ	0.118	0.188	0.125	0.118
ทุเรียน	1	0.75	0.857	1
มังคุด	0.083	0.182	0.091	0.083
ลำไย	0.1	0.1	0	0.1
ส้มโอ	0	0.091	0	0
ส้มเขียวหวาน	0.286	0.385	0.214	0.286
มะม่วง	0	0.091	0	0

ตารางที่ 4.7 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่ที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “มังคุด”

	พิมพ์ถูก มังคุด	พิมพ์ผิด มัวคุด	พิมพ์ตก มังคุ	พิมพ์เกิน มังงคุด
ไม้ผลเศรษฐกิจ	0.059	0.059	0.063	0.059
ทุเรียน	0.083	0.083	0.091	0.083
มังคุด	1	0.714	0.833	1
ลำไย	0	0	0	0
ส้มโอ	0.1	0.1	0.111	0.1
ส้มเขียวหวาน	0.063	0.133	0.067	0.063
มะม่วง	0.222	0.222	0.25	0.222

ตารางที่ 4.8 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่ที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า “ลำไย”

	พิมพ์ถูก ลำไย	พิมพ์ผิด ลำใย	พิมพ์ตก ลำย	พิมพ์เกิน ลำไยบ
ไม้ผลเศรษฐกิจ	0.143	0.067	0.071	0.133
ทุเรียน	0.1	0.1	0.111	0.091
มังคุด	0	0	0	0
ลำไย	1	0.6	0.75	0.8
ส้มโอ	0	0	0	0
ส้มเขียวหวาน	0.071	0.071	0.077	0.067
มะม่วง	0	0	0	0

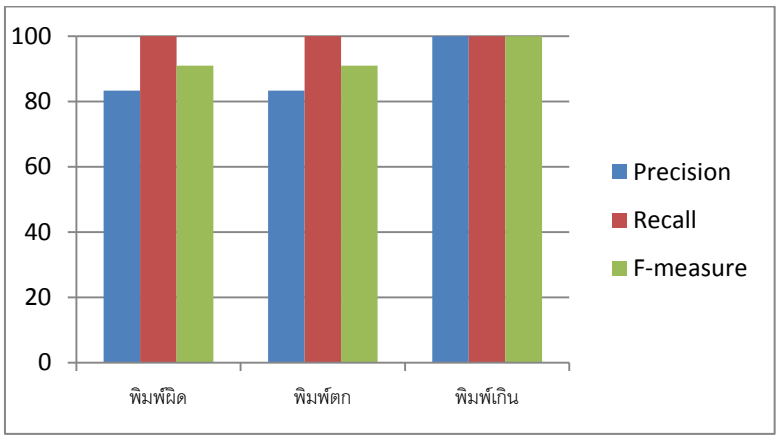
ตารางที่ 4.9 ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดที่ที่มีผลต่อคำที่มีข้อผิดพลาด ตัวอย่างคำว่า ”ส้มเขียวหวาน”

	พิมพ์ถูก ส้มเขียวหวาน	พิมพ์ผิด ส้มเชียวหวาย	พิมพ์ตก ส้มเขียวหวน	พิมพ์เกิน ส้มเขียววหวาน
ไม้ผลเศรษฐกิจ	0.15	0.158	0.158	0.15
ทุเรียน	0.286	0.214	0.308	0.286
มังคุด	0.063	0.067	0.067	0.063
ลำไย	0.071	0.077	0.077	0.072
ส้มโอ	0.231	0.25	0.25	0.231
ส้มเขียวหวาน	1	0.75	0.909	1
มะม่วง	0.143	0.154	0.154	0.143

จากผลการทดลองแสดงให้เห็นว่าค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดจะมีค่าอยู่ระหว่าง 0 ถึง 1 โดยค่า 0 แสดงว่าไม่มีความคล้ายคลึงกันเลย และค่า 1 แสดงว่ามีความคล้ายคลึงกันมากที่สุด จากการวัดประสิทธิภาพด้วยการใช้ค่าความแม่นยำ ค่าการจำได้ และอัตราการรู้จำซึ่งคำนวณได้จากการวิเคราะห์ค่าสัมประสิทธิความคล้าย โดยหากค่าความคล้ายตั้งแต่ 0.5 แปลว่าคำนั้นถูกเลือก นอกจากนั้นจะถูกตีความว่าไม่ถูกเลือก และข้อสังเกตพบว่า การวัดประสิทธิภาพความคล้ายด้วยสัมประสิทธิ์แจ๊คการ์ด สามารถทำนายค่าได้ถูกต้องดี และมีเสถียรภาพในการทำงานสูงมาก แต่ในกรณีที่มีการพิมพ์ตก และพิมพ์เกิน หรือมีการพิมพ์ซ้ำตัวเดิม จะยังคงให้ผลลัพธ์ที่มีค่าสูง นั่นหมายความว่า การวัดความคล้ายด้วยสัมประสิทธิ์ความคล้ายของแจ็คการ์ดไม่สามารถตรวจสอบการมีอยู่ของตัวอักษรที่ซ้ำกันได้ ตัวอย่างเช่น คำว่า “ไม้ผลเศรษฐกิจ” พิมพ์ตกเป็น “ไม้ผลเศรษฐกจ” และพิมพ์เกินหรือพิมพ์ซ้ำเป็น “ไม้ผลเศรรษฐกิจ” มีค่ากับ “0.846” และ “0.923” ตามลำดับ ซึ่งผลลัพธ์ความคล้ายมีค่าสูง เป็นต้น ผลปรากฏดังตารางที่ 4.10 และแสดงดังภาพที่ 4.5

ตารางที่ 4.10 แสดงการวัดประสิทธิภาพค่าสัมประสิทธิ์ความคล้ายตั้งแต่ 0.5

Keyword	Precision	Recall	F-Measure
พิมพ์ผิด หรือสะกดผิด	83.33	100	90.90
พิมพ์ตก	83.33	100	90.90
พิมพ์เกิน	100	100	100



ภาพที่ 4.5 แผนภูมิแสดงการวัดประสิทธิภาพค่าสัมประสิทธิ์ความคล้ายตั้งแต่ 0.5

4.3 การค้นหาเชิงความหมาย

ในส่วนการสืบค้นเชิงความหมายจะอธิบายการวัดความคล้ายเชิงความหมายระหว่างคำค้นกับเอกสาร โดยการให้ค่าความสำคัญจากการคำนวณจากค่าระยะทางสั้นที่สุดของรายการคำที่เป็นโครงสร้างซึ่งเก็บอยู่ในคลังคำ ซึ่งผู้วิจัยเสนอมี 3 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การกำหนดค่าความสำคัญของคำค้น

การค้นคืนเอกสารในบางครั้งผู้ใช้งานใช้คำค้นมากกว่าหนึ่งคำเพื่อให้สามารถค้นหาเอกสารที่ต้องการได้อย่างแม่นยำมากยิ่งขึ้น แต่เมื่อเอกสารที่เป็นผลลัพธ์ของการสืบค้นประกอบด้วยคำค้นในจำนวนที่เท่ากันจึงเป็นการยากที่จะลำดับความสำคัญของเอกสารที่เป็นผลลัพธ์นั้น ดังตัวอย่าง

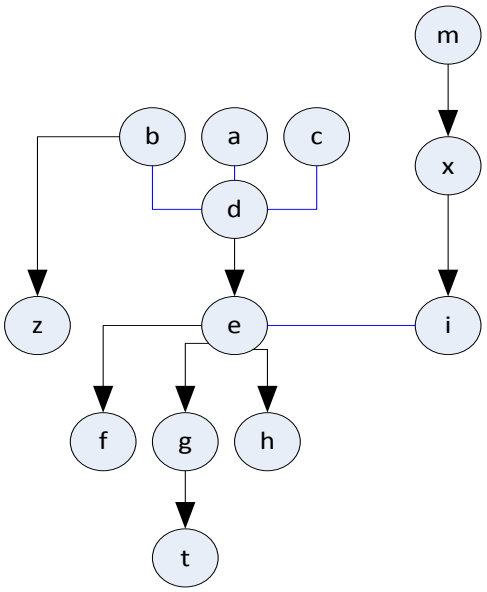
คำค้น : ทุเรียน , พืชส่งออก

เอกสารที่ 1 : ทุเรียน , ไม้ผล , ปุ๋ย

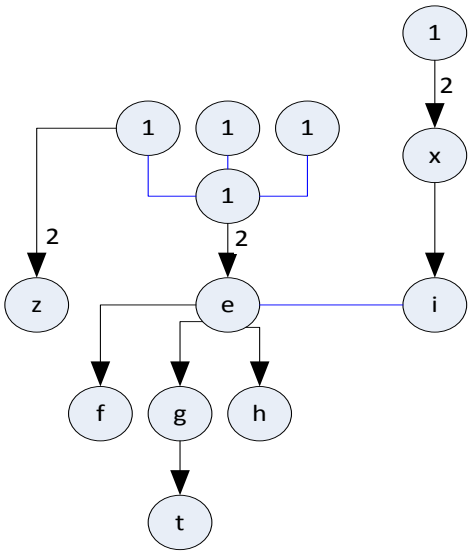
เอกสารที่ 2 : มังคุด , พืชส่งออก , ไม้ผล

ด้วยเหตุนี้ผู้วิจัยจึงได้ออกแบบการกำหนดความสำคัญของคำค้นซึ่งอาจสามารถช่วยให้สามารถจัดลำดับเอกสารได้ตรงตามความต้องการของผู้ใช้มากยิ่งขึ้น โดยการกำหนดความสำคัญให้กับคำตามโครงสร้างความสัมพันธ์โดยมีลำดับความสำคัญที่แตกต่างกันตามลำดับความลึก เพื่อเป็นการอธิบายขั้นตอนวิธีในการค้นคืน ผู้วิจัยจะได้อธิบายในลำดับต่อไปโดยกำหนดตัวอย่างให้การค้นคืนสำหรับคำต่าง ๆ ที่ได้จัดเก็บไว้แทนด้วยตัวอักษร a-z และแสดงโครงสร้างตัวอย่าง ดังการดำเนินการค้นหาตามขั้นตอนวิธีที่ได้กำหนดขึ้นมาตามตัวอย่างที่ได้แสดงดังต่อไปนี้

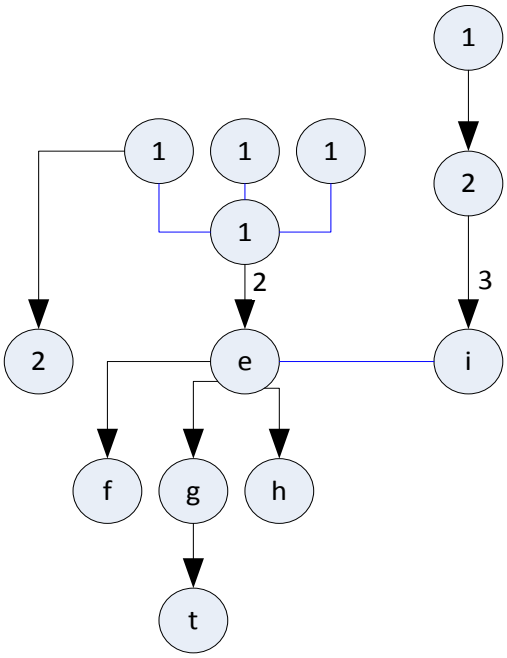
1. ตัวอย่างแสดงโครงสร้างความสำพันธ์ของคำสำคัญที่เก็บอยู่ในคลังคำ



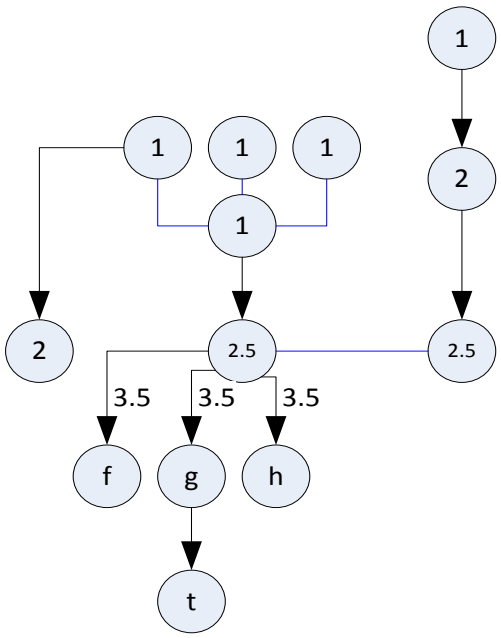
2. กำหนดให้คำที่ไม่มีคำกว้างกว่ามีค่าความลึกเป็น 1 และ กำหนดให้คำที่เหมือนกันหรือเกี่ยวข้องมีค่าเท่ากัน แสดงดังตัวอย่าง



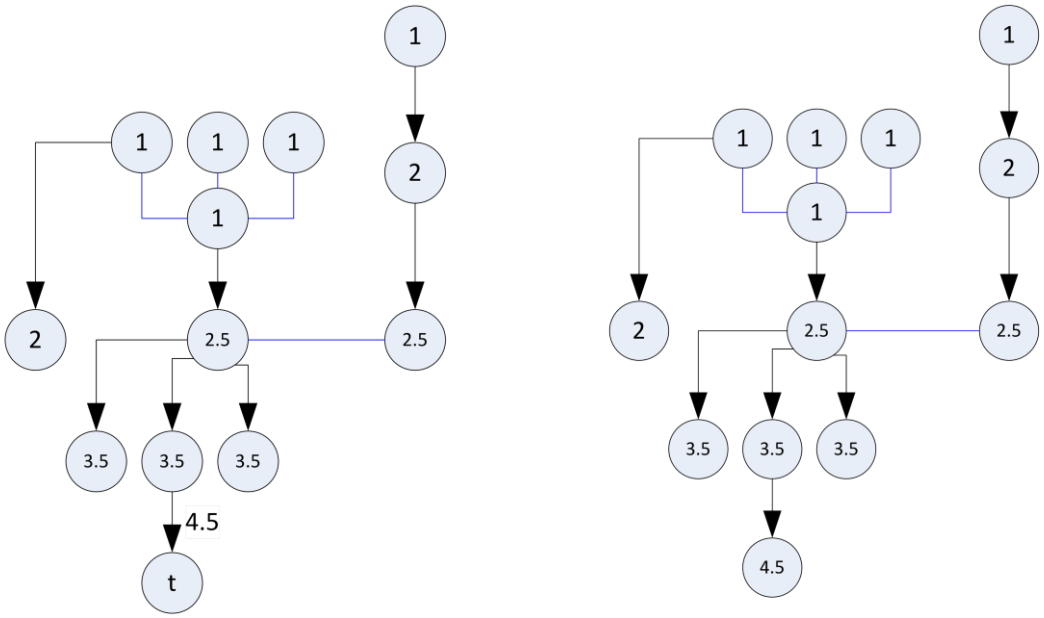
3. กำหนดค่าคาดหวังให้คำแคบกว่ามีค่าเพิ่มหนึ่ง แสดงดังตัวอย่าง



4. คำนวณค่าความสำคัญจากผลรวมค่าคาดหวังหารด้วยจำนวนความคาดหวัง แสดงดังตัวอย่าง



5. ทำซ้ำข้อ 2 ถึง 4 จนครบทั้งโครงสร้าง แสดงดังตัวอย่าง



ขั้นตอนที่ 2 คำนวณหาค่าน้ำหนักของคำค้น

ในขั้นตอนนี้จะทำการปรับค่าน้ำหนักของคำค้นให้สามารถนำมาวิเคราะห์ความสำคัญได้ โดยการปรับค่าความสำคัญให้ได้ผลรวมของค่าน้ำหนักเท่ากับหนึ่ง ดังสมการ 4.1

$$Q_i = \frac{\mathrm{L}_i}{\sum L} \qquad (4.1)$$

Q_i แทน ค่าน้ำหนักของคำค้นที่ i

L_i แทน ค่าความสำคัญของคำค้นที่ i

$\sum L$ แทน ผลรวมค่าความสำคัญของคำค้นทั้งหมด

ตารางที่ 4.11 แสดงตัวอย่างการคำนวณหาค่าน้ำหนักของคำค้น

	q1	q2 q3	q1 q4 q5
ค่าความสำคัญ	-	2 1	1 2 3
น้ำหนักของคำค้น	1	0.67 0.33	0.17 0.33 0.5

ในที่นี้ กำหนดให้ q1 q2 q3 q4 และ q5 หมายถึงคำค้นจากผู้ใช้

สดมภ์ที่ 2 q1 เป็นคำค้นเดียวจึงมีค่าน้ำหนักเป็น 1

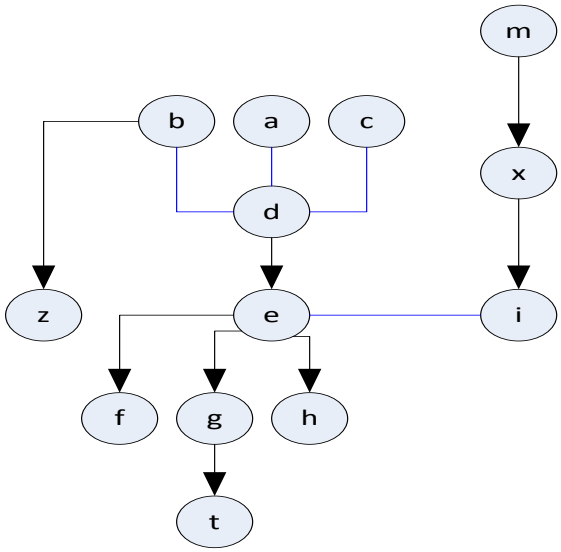
สดมภ์ที่ 3 q2 มีค่าความสำคัญ 2 และ q3 มีค่าความสำคัญ 1 ทำการปรับค่าน้ำหนักเป็น 0.67 และ 0.33 ตามลำดับ

สดมภ์ที่ 4 q1 q4 และ q5 มีค่าความสำคัญ 1 2 และ 3 ทำการปรับค่าน้ำหนักเป็น 0.17 0.33 และ 0.5 ตามลำดับ

ขั้นตอนที่ 3 การหาเส้นทางสั้นที่สุดตามโครงสร้าง

การสืบค้นเอกสารเชิงความหมายจำเป็นต้องอาศัยโครงสร้างความสัมพันธ์เพื่ออธิบายความเกี่ยวข้องกันเชิงความหมายของคำสืบค้นกับเอกสาร ซึ่งความสัมพันธ์ของคำค้นกับคำในเอกสารสามารถคำนวณหาระยะทางได้โดยการหาวัดระยะทางของความสัมพันธ์ระหว่างคำโดยการอ้างอิงเส้นทางตามโครงสร้างความสัมพันธ์ มีขั้นตอนวิธีแสดงดังตัวอย่าง ซึ่งมีผลลัพธ์แสดงดังตารางที่ 4.12

1. ตัวอย่างแสดงโครงสร้างความสัมพันธ์ของคำสำคัญที่เก็บอยู่ในคลังคำ

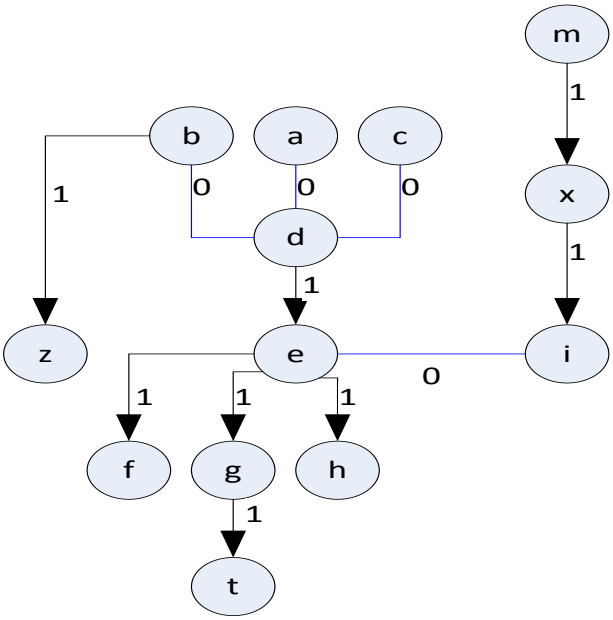


2. วิธีการคำนวณแสดงดังตัวอย่าง ดังนี้

2.1 กำหนดให้คำที่เหมือนกันหรือเกี่ยวข้องมีค่าระยะทางเป็น 0

2.2 กำหนคำที่แคบกว่าหรือกว้างกว่ามีค่าระยะทางเป็น 1

2.3 คำนวณค่าระยะทางโดยการหาผลรวมในเส้นทางที่สั้นที่สุดจากจุดหนึ่งไปยังอีกจุด



ตารางที่ 4.12 แสดงผลลัพธ์การคำนวณหาเส้นทางสั้นที่สุดตามโครงสร้าง

	a	b	c	d	e	f	g	h	i	m	t	x	z
a	0	0	0	0	1	2	2	2	1	3	3	2	1
b	0	0	0	0	1	2	2	2	1	3	3	2	1
c	0	0	0	0	1	2	2	2	1	3	3	2	1
d	0	0	0	0	1	2	2	2	1	3	3	2	1
e	1	1	1	1	0	1	1	1	0	2	2	1	2
f	2	2	2	2	1	0	2	2	1	3	3	2	3
g	2	2	2	2	1	2	0	2	1	3	1	2	3
h	2	2	2	2	1	2	2	0	1	3	3	2	3
i	1	1	1	1	0	1	1	1	0	2	2	1	2
m	3	3	3	3	2	3	3	3	2	0	4	1	4
t	3	3	3	3	2	3	1	3	2	4	0	3	4
x	2	2	2	2	1	2	2	2	1	1	3	0	3
z	1	1	1	1	2	3	3	3	2	4	4	3	0

4.4 การแสดงผลลัพธ์

การแสดงผลลัพธ์จะเป็นขั้นตอนต่อจากการค้นหาเชิงความหมาย หลังจากที่ได้กลุ่มคำค้นและรายการเอกสารความรู้ที่เกี่ยวข้องมาแล้ว จะทำการเปรียบเทียบความคล้ายคลึงของคำค้นกับเอกสารแต่ละรายการ และเรียงลำดับค่าความคล้ายคลึงจากมากไปหาน้อยเพื่อแสดงผลลัพธ์ในการค้นหา ดังนี้

ขั้นตอนที่ 1 การหาค่าน้ำหนักของเอกสารต่อคำค้น

ในขั้นตอนนี้จะทำการปรับค่าน้ำหนักของเอกสารต่อคำค้นโดยนำผลลัพธ์การคำนวณหาเส้นทางสั้นที่สุดตามโครงสร้างมาทำการคำนวณหาค่าน้ำหนักของเอกสารต่อคำค้นให้สามารถนำมาวิเคราะห์ความคล้ายคลึงเชิงความหมายได้ โดยการปรับให้สอดคล้องกับระยะทางระหว่างคำในเอกสารกับคำค้นที่ป้อนโดยผู้ใช้ ดังสมการ 4.2

$$D_k = \frac{\sum \exp(-dis(d_j,q_i))}{N_k} \qquad (4.2)$$

D_k แทน ค่าน้ำหนักของเอกสารต่อคำค้นที่ k

$dis(d_j, q_i)$ แทน ค่าระยะทางระหว่างคำในเอกสาร d_j กับคำค้น q_i

N_k แทน จำนวนคำในเอกสาร

ตารางที่ 4.13 แสดงตัวอย่างการคำนวณหาน้ำหนักของเอกสารต่อคำค้น

	q1	q2	q3	q1	q4	q5
kDoc1	0	4	5	0	1	2
kDoc2	3	5	6	3	2	3
kDoc3	4	0	1	4	3	4
น้ำหนักเอกสารต่อคำค้น	0.356	0.342	0.126	0.356	0.184	0.068

ในที่นี้ กำหนดให้ kDoc1 kDoc2 และ kDoc3 หมายถึงคำสำคัญที่ปรากฏในเอกสาร

สดมภ์ที่ 2 บรรทัดที่ 2 แสดงค่าระยะทางสั้นที่สุดระหว่างคำสำคัญ (kDoc1) ที่ปรากฏในเอกสารกับ

คำค้น (q1) จากผู้ใช้

สดมภ์ที่ 2 บรรทัดที่ 3 ... (kDoc2) ... (q1) ...

สดมภ์ที่ 2 บรรทัดที่ 4 ... (kDoc3) ... (q1) ...

สดมภ์ที่ 2 บรรทัดที่ 5 แสดงค่าน้ำหนักเอกสารต่อคำค้นด้วยการคำนวณตามสมการที่ 4.2

<u>ขั้นตอนที่ 2</u> การหาคะแนนความคล้ายเชิงความหมายของเอกสาร

ในขั้นตอนนี้จะทำการคำนวณคะแนนความคล้ายเชิงความหมายของเอกสารเพื่อใช้ในการลำดับความสำคัญของเอกสารให้ตรงตามความต้องการของผู้ใช้ได้มากที่สุด โดยใช้ค่าน้ำหนักของคำค้นในขั้นตอนที่ 2 และใช้ค่าน้ำหนักของเอกสารต่อคำค้น ในขั้นตอนที่ 4 มาคำนวณ ดังสมการ 4.3

$$S_k = \sum(Q_i * D_k) \qquad (4.3)$$

S_k แทน คะแนนความคล้ายเชิงความหมายของเอกสารที่ k

Q_i แทน ค่าน้ำหนักของคำค้นที่ i

D_k แทน ค่าน้ำหนักของเอกสารต่อคำค้นที่ k

ตารางที่ 4.14 แสดงตัวอย่างการคำนวณหาคะแนนความคล้ายเชิงความหมายของเอกสาร

	q1	q2	q3	q1	q4	q5
น้ำหนักคำค้น	1	0.67	0.33	0.17	0.33	0.5
คะแนนเอกสารต่อคำค้น	0.356	0.342	0.126	0.356	0.184	0.068
คะแนนเอกสาร	0.356	0.27		0.155		

สดมภ์ที่ 2 บรรทัดที่ 2 แสดงค่าน้ำหนักคำค้นจากผู้ใช้

สดมภ์ที่ 2 บรรทัดที่ 3 แสดงค่าน้ำหนักเอกสารต่อคำค้น

สดมภ์ที่ 2 บรรทัดที่ 4 แสดงผลการคำนวณคะแนนความคล้ายเชิงความหมายของเอกสารตามสมการ 4.3

- **ผลการวิเคราะห์การทดสอบความคล้ายเชิงความหมาย**

ผลการวิเคราะห์การทดสอบการคำนวณหาคะแนนความคล้ายเชิงความหมายสำหรับเอกสารความรู้ แสดงดังตัวอย่างปรากฏในตารางที่ 4.15 ซึ่งตัวอย่างแสดงผลการคำนวณค่าคะแนนความคล้ายเชิงความหมายของเอกสาร กำหนดให้ตัวอักษร a ถึง p แทน คำสำคัญสำหรับค้นหา เช่น ตัวอักษร f แทน 1 คำค้น o d แทน 2 คำค้น และ a d n แทน 3 คำค้น เป็นต้น และกำหนดให้แต่ละเอกสาร (Document) มีคำสำคัญ เช่น เอกสารที่ 1 ประกอบด้วยคำสำคัญ a e k เอกสารที่ 2 ประกอบด้วยคำสำคัญ b e k เป็นต้น จากการทดสอบพบว่า เช่น ผลการคำนวณคำค้น f กับเอกสารที่ 1 ได้ค่าคะแนนความคล้ายเชิงความหมายเท่ากับ 0.02 คำค้น o และ d กับเอกสารที่ 1 ได้ค่าคะแนนความคล้ายเชิงความหมายเท่ากับ 0.11 เป็นต้น

ตารางที่ 4.15 ตัวอย่างแสดงผลการคำนวณหาคะแนนความคล้ายเชิงความหมายของเอกสาร

Document		การสืบค้นเชิงความหมาย					
		f	g	o d	b e	a d n	f h p
1	a e k	0.02	0.13	0.11	0.29	0.15	0.03
2	b e k	0.02	0.13	0.11	0.46	0.11	0.03
3	b f k	0.35	0.17	0.09	0.20	0.08	0.15
4	c f k	0.36	0.21	0.01	0.00	0.01	0.15
5	c f j	0.36	0.21	0.00	0.00	0.00	0.15
6	c g j	0.07	0.50	0.01	0.00	0.01	0.11
7	d g j m	0.05	0.36	0.22	0.07	0.20	0.08
8	d g i l	0.06	0.44	0.16	0.05	0.15	0.09
9	d g i	0.06	0.46	0.21	0.07	0.20	0.10
10	d g h	0.17	0.46	0.21	0.07	0.20	0.27

- **ผลการวิเคราะห์การทดสอบการเลือกเอกสาร**

ผลการวิเคราะห์การทดสอบการเลือกเอกสารการสืบค้นแบบการจับคู่กับ ระบุโดยผู้เชี่ยวชาญ และการสืบค้นเชิงความหมายกับระบุโดยผู้เชี่ยวชาญ แสดงดังตารางที่ 4.16 และตารางที่ 4.17

ตารางที่ 4.16 แสดงผลเอกสารที่ถูกเลือกระหว่างวิธีการสืบค้นแบบจับคู่กับระบุโดยผู้เชี่ยวชาญ

Document		การสืบค้นแบบจับคู่						ระบุโดยผู้เชี่ยวชาญ					
		f	g	o d	b e	a d n	f h p	f	g	o d	b e	a d n	f h p
1	a e k				*	*			*	*	*	*	
2	b e k				*				*		*		
3	b f k	*			*		*	*	*		*		*
4	c f k	*					*	*	*				*
5	c f j	*					*	*	*				*
6	c g j		*						*				*
7	d g j m		*	*		*			*	*	*	*	*
8	d g i l		*	*		*			*	*	*	*	*
9	d g i		*	*		*			*	*	*	*	*
10	d g h		*	*		*	*		*	*	*	*	*

จากตารางที่ 4.16 แสดงผลการเลือกเอกสารโดยการสืบค้นแบบจับคู่ (matching) และการเลือกเอกสารโดยผู้เชี่ยวชาญ โดยกำหนดให้ * แทนเอกสารที่ถูกเลือก ซึ่งหมายความว่าเอกสารนั้นมีความเกี่ยวข้องกับคำค้นที่ผู้ใช้ใส่เข้ามา 1 คำค้น f และ g 2 คำค้น o d และ b e 3 คำค้น a d n และ f h p ผลการทดสอบพบว่า การสืบค้นแบบจับคู่จะมีความถูกต้องกับเอกสารในกรณีที่เอกสารนั้นมีคำตรงกับคำที่ผู้ใช้ระบุเข้ามาโดยไม่สามารถหาคำกว้างกว่า (Broader Term-BT) คำแคบกว่า (Narrower Term-NT) และ คำเกี่ยวข้อง (Related Term-RT) เมื่อเทียบกับคำที่ผู้เชี่ยวชาญระบุ เช่น ในเอกสาร (Document) ที่ 1 ประกอบด้วยคำสำคัญ a e และ k การสืบค้นแบบจับคู่ต้องระบุคำค้นที่เข้ามาให้ตรงกับคำสำคัญที่มีอยู่ในเอกสารเท่านั้น แต่ความคาดหวังของผู้เชี่ยวชาญหรือผู้เชี่ยวชาญกำหนดในเอกสาร ที่ 1 ควรจะสืบค้นจากคำ g และ o d เข้ามาควรจะพบเอกสารที่ 1 ได้ด้วย

ตารางที่ 4.17 แสดงผลเอกสารที่ถูกเลือกระหว่างการสืบค้นเชิงความหมายกับระบุโดยผู้เชี่ยวชาญ

		การสืบค้นเชิงความหมาย						ระบุโดยผู้เชี่ยวชาญ					
Document		f	g	o d	b e	a d n	f h p	f	g	o d	b e	a d n	f h p
1	a e k		*		*	*			*	*	*	*	
2	b e k		*		*				*		*		
3	b f k	*	*		*		*	*	*		*		*
4	c f k	*	*				*	*	*				*
5	c f j	*	*				*	*	*				*
6	c g j		*						*				*
7	d g j m		*	*		*			*	*	*	*	*
8	d g i l		*	*		*			*	*	*	*	*
9	d g i		*	*		*			*	*	*	*	*
10	d g h	*	*	*		*	*		*	*	*	*	*

จากตารางที่ 4.17 แสดงผลการเลือกเอกสารโดยการสืบค้นเชิงความหมาย และการเลือกเอกสารโดยผู้เชี่ยวชาญ โดย * แทนเอกสารที่มีค่าความคล้ายเชิงความหมายสูงกว่า 0.125 เอกสารนั้นจะถูกเลือก ซึ่งหมายความว่าเอกสารนั้นมีความเกี่ยวข้องกับคำค้นที่ผู้ใช้ใส่เข้ามา ผลการทดสอบการสืบค้นเชิงความหมายสามารถค้นหาคำหลัก คำพ้อง คำกว้างกว่า (Broader Term-BT) คำแคบกว่า (Narrower Term-NT) และ คำเกี่ยวข้อง (Related Term-RT) เมื่อเทียบกับคำที่ผู้เชี่ยวชาญระบุ

4.5 ผลการประเมินระบบ

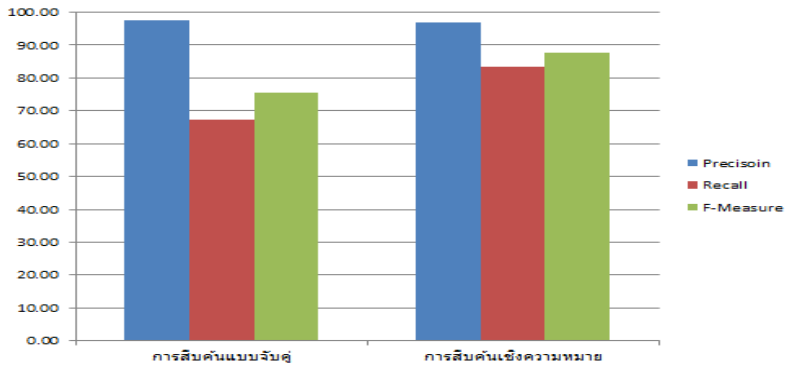
ผลการประเมินระบบประกอบด้วยสองส่วน ได้แก่ ผลการประเมินความถูกต้องของระบบ และผลการประเมินความสามารถของระบบ โดยมีรายละเอียดในแต่ละส่วน ดังนี้

- **ผลการประเมินความถูกต้องของระบบ**

ผลการวิเคราะห์การทดสอบประสิทธิภาพการเปรียบเทียบระหว่างการสืบค้นแบบจับคู่กับการสืบค้นเชิงความหมาย แสดงดังตารางที่ 4.18

ตารางที่ 4.18 แสดงการวัดประสิทธิภาพผลการเปรียบเทียบระหว่างการสืบค้นแบบจับคู่กับการสืบค้นเชิงความหมาย

Method	Precision	Recall	F-Measure
การสืบค้นแบบจับคู่	100.00	55.77	67.31
การสืบค้นเชิงความหมาย	96.84	83.54	87.55



ภาพที่ 4.6 แผนภูมิสรุปผลการประเมินความถูกต้องของการสืบค้น

จากตารางที่ 4.18 และภาพที่ 4.6 พบว่า ค่าความแม่นยำ (Precision) ค่าความถูกต้อง (Recall) และค่าอัตราการรู้จำ (F-measure) มีความแตกต่างกันในแต่ละวิธี สำหรับค่าความแม่นยำ ของการสืบค้นแบบจับคู่มีค่าเท่ากับ 100 เปอร์เซ็นต์ หมายความว่าทุกเอกสารที่เลือกมามีความเกี่ยวข้องทั้งสิ้น แต่การสืบค้นแบบเชิงความหมายมีค่าความแม่นยำ เท่ากับ 96.84 เปอร์เซ็นต์ ซึ่งหมายความว่าการเลือกในบางครั้งมีเอกสารที่ไม่เกี่ยวข้องติดไปด้วย

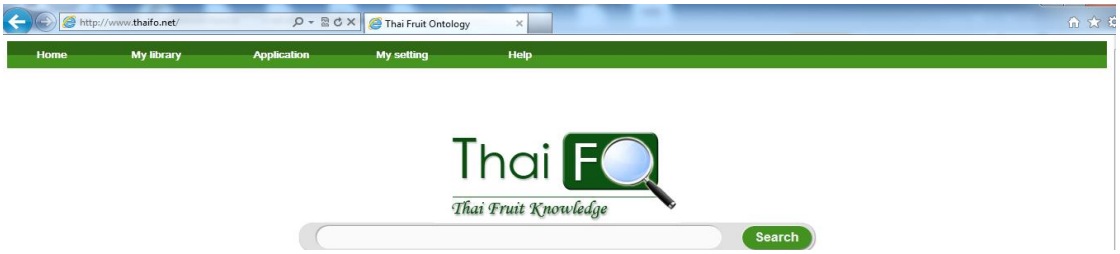
สำหรับค่าความถูกต้อง ของการสืบค้นแบบจับคู่มีค่าเท่ากับ 55.77 เปอร์เซ็นต์ หมายความว่ามีความสามารถสืบค้นเอกสารได้ตรงตามความต้องการไม่ครบถ้วน แต่การสืบค้นแบบเชิงความหมายมีค่าเท่ากับ 83.54 เปอร์เซ็นต์ ซึ่งหมายความว่ามีความสามารถสืบค้นเอกสารได้ตรงตามความต้องการครบถ้วนมากกว่า และเมื่อพิจารณาค่าความระลึก ซึ่งเป็นการหาค่าเฉลี่ยที่ให้ความสำคัญ

กับความแม่นยำและความครบถ้วนเท่า ๆ กันพบว่าการสืบค้นเชิงความหมายสามารถค้นหาเอกสารที่เกี่ยวข้องได้ตรงความต้องการและมีความครบถ้วนดีกว่าการสืบค้นแบบจับคู่ธรรมดาอย่างเห็นได้ชัด นั้นเป็นเพราะการสืบค้นเชิงความหมายสามารถค้นหาหลัก คำพ้อง คำกว้างกว่า (Broader Term-BT) คำแคบกว่า (Narrower Term-NT) และ คำเกี่ยวข้อง (Related Term-RT) ได้ ซึ่งการสืบค้นแบบจับคู่จะมีความถูกต้องกับเอกสารในกรณีที่เอกสารนั้นมีคำตรงกับคำที่ผู้ใช้ระบุเข้ามาเท่านั้น

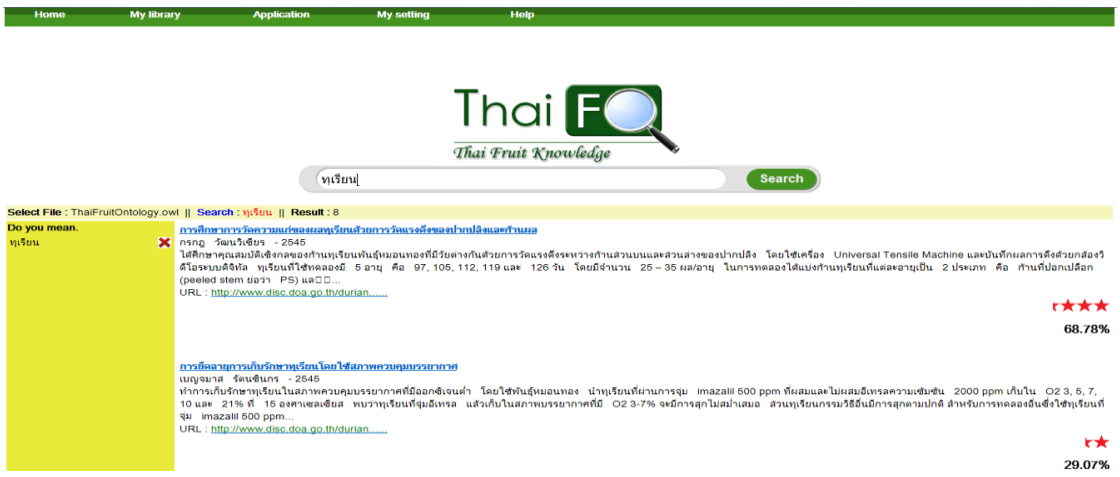
- **การประเมินความสามารถของระบบ**

การทดสอบสืบค้นเชิงความหมายผ่านส่วนติดต่อกับผู้ใช้งานบนเว็บบราวเซอร์จากฐานความรู้ที่สร้างขึ้น ซึ่งผู้วิจัยได้ทำการ Upload ผ่านเครือข่ายอินเทอร์เน็ต โดยผู้ใช้งานสามารถเข้าไปสืบค้นข้อมูลได้ที่ http://www.thaifo.net และมีผลการประเมินความสามารถของระบบ ดังนี้

1. ตัวอย่างผลการสืบค้นจากผู้เชี่ยวชาญที่ประเมินความสามารถของระบบ



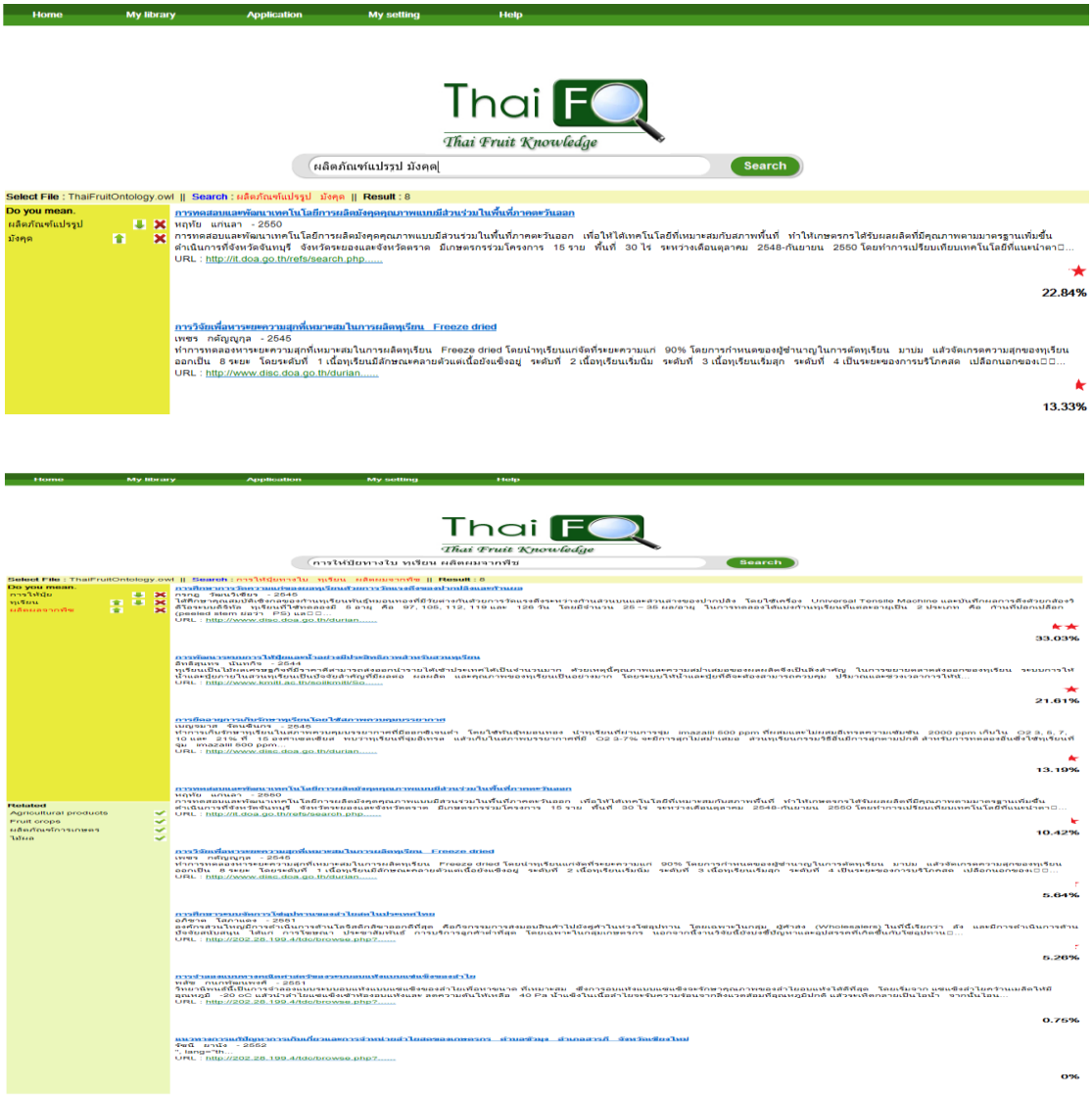
ภาพที่ 4.7 แสดงหน้าต่างหลักของเว็บไซต์การสืบค้นเชิงความหมาย



ภาพที่ 4.8 แสดงตัวอย่างผลลัพธ์จากการสืบค้นเชิงความหมายกรณี 1 คำค้น

ภาพที่ 4.8 แสดงผลลัพธ์จากการสืบค้นเชิงความหมายกรณี 1 คำค้น ซึ่งมีการพิมพ์ผิดจากคำว่า ‘Durian’ เป็นคำว่า ‘Durin’ พบว่าระบบสามารถหาความคล้ายคลึงเป็นคำ ‘Durian’ ได้โดยอัตโนมัติและแสดงรายละเอียดเอกสารที่มีความเกี่ยวข้องกับคำค้น Durian โดยระบบได้มีระบบการ

ป้องการพิมพ์ผิดจากการใช้สัมประสิทธิ์ความคล้ายของแจ๊คการ์ดเพื่อค้นหาคำที่ถูกต้องในการค้นหาเอกสารได้อย่างครบถ้วนแม่นยำ



ภาพที่ 4.9 แสดงตัวอย่างผลลัพธ์จากการสืบค้นเชิงความหมาย 2 และ 3 คำค้น

ภาพที่ 4.9 แสดงผลลัพธ์จากการสืบค้นเชิงความหมาย 2 คำค้น ซึ่งสืบค้นโดยใช้คำว่า ‘ผลิตภัณฑ์แปรรูป’ กับคำว่า ‘มังคุด’ และ 3 คำค้น ซึ่งสืบค้นโดยใช้คำว่า ‘การให้ปุ๋ยทางใบ ทุเรียน ผลิตผมจากพืช’ และมีการแสดงรายละเอียดเอกสารที่มีความเกี่ยวข้องกับคำค้น ‘ผลิตภัณฑ์แปรรูป มังคุด’ และ ‘การให้ปุ๋ยทางใบ ทุเรียน ผลิตผมจากพืช’ ซึ่งระบบจะทำการจัดเรียงลำดับตามค่าของคะแนนซึ่งได้มาจากการคำนวณหาคะแนนความคล้ายเชิงความหมายของเอกสารที่คิดออกมาเป็นเปอร์เซ็นต์

ข้อสังเกตพบว่า สำหรับการสืบค้นเชิงความหมายกับเอกสารความรู้ด้านการเกษตรโดยใช้คำค้นเดียวนั้นค่าคะแนนจะขึ้นกับค่าระยะทางสั้นที่สุดระหว่างคำค้นกับคำในเอกสารจากนั้นทำการเฉลี่ยด้วยจำนวนคำในเอกสาร วิธีการนี้จะส่งผลให้ค่าคะแนนเอกสารสูงในกรณีที่คำในเอกสารมีน้อยและมีระยะทางสั้น ส่วนการสืบค้นเชิงความหมายโดยใช้คำค้นหลายคำผลการสืบค้นจะถูกปรับโดยใช้ค่าความสำคัญของคำค้นแต่ละคำส่งผลให้ค่าคะแนนเอกสารต่ำลงแต่ความถูกต้องของเอกสารตามความหมายโดยรวมจะสูงขึ้น เช่น การสืบค้นโดยใช้คำว่า ‘ทุเรียน’ ผลลัพธ์คือ เอกสารเกี่ยวกับทุเรียน ผลไม้ การสืบค้นโดยใช้คำว่า ‘ทุเรียน แปรรูป’ ผลลัพธ์คือ เอกสารเกี่ยวกับ ทุเรียน ทุเรียนกวน ทุเรียนทอด เป็นต้น

2. ผลการประเมินความสามารถระบบของผู้เชี่ยวชาญ

2.1 ข้อมูลทั่วไปของผู้ตอบแบบประเมินความสามารถระบบสืบค้น ดังตารางที่ 4.19

ตารางที่ 4.19 ข้อมูลทั่วไปของผู้ตอบแบบประเมินความสามารถระบบสืบค้น

ข้อมูลทั่วไปของผู้ตอบแบบประเมิน	จำนวน	ร้อยละ
1. เพศ		
ชาย	9	60.00
หญิง	6	40.00
รวม	15	100.00
2. ระดับการศึกษา		
ต่ำกว่าปริญญาตรี	-	-
ปริญญาตรีหรือเทียบเท่า	5	33.30
ปริญญาโท	7	46.70
ปริญญาเอก	3	20.00
รวม	15	100.00

จาก

ตารางที่ 4.19 พบว่า ผู้ตอบแบบสอบถามประเมินความสามารถระบบสืบค้นส่วนใหญ่เป็นเพศชาย (ร้อยละ 60.00) ระดับการศึกษาปริญญาโท (ร้อยละ 46.7) รองลงมาปริญญาตรีหรือเทียบเท่า (ร้อยละ 33.30)

2.2 ความคิดเห็นเกี่ยวกับความสามารถของระบบสืบค้น

ตารางที่ 4.20 ความคิดเห็นเกี่ยวกับความสามารถของระบบสืบค้นโดยรวมในแต่ละด้าน

ความสามารถของระบบสืบค้น	$\overline{X}$	S.D.	ระดับความคิดเห็น
ด้านการประเมินความสามารถของระบบ (Functional requirement test)	4.18	0.32	ดี
ด้านการประเมินความถูกต้องของระบบ (Functional test)	4.05	0.61	ดี
ด้านการประเมินความง่ายในการใช้งาน (Usability test)	4.43	0.19	ดี
ด้านการประเมินประสิทธิภาพของระบบ (Performance test)	3.98	0.43	ดี
ด้านการประเมินด้านการรักษาความปลอดภัย (Security test)	3.38	0.49	ปานกลาง
โดยรวม	4.00	0.14	ดี

จาก

ตารางที่ 4.20 พบว่า ผู้เชี่ยวชาญมีความคิดเห็นด้วยเกี่ยวกับความสามารถของระบบสืบค้นโดยรวมอยู่ในระดับดี ($\overline{X}$ = 4.00) เมื่อพิจารณาเป็นรายด้านอยู่ในระดับดี จำนวน 4 ด้าน โดยเรียงลำดับค่าเฉลี่ยจากมากไปหาน้อย ดังนี้ ด้านการประเมินความง่ายในการใช้งาน ($\overline{X}$ = 4.43) ด้านการประเมินความสามารถของระบบ ($\overline{X}$ = 4.18) ด้านการประเมินความถูกต้องของระบบ ($\overline{X}$ = 4.05) ด้านการประเมินประสิทธิภาพของระบบ ($\overline{X}$ = 3.98) และอยู่ในระดับปานกลาง จำนวน 1 ข้อ ได้แก่ ด้านการประเมินด้านการรักษาความปลอดภัย ($\overline{X}$ = 3.38)

ตารางที่ 4.21 ความคิดเห็นเกี่ยวกับด้านการประเมินความสามารถของระบบ

ความสามารถของระบบสืบค้น ด้านการประเมินความสามารถของระบบ	$\overline{X}$	S.D.	ระดับความคิดเห็น
1. ความสามารถของส่วนนำเข้าข้อคำถามตรงตามความต้องการของท่านในระดับใด	4.60	0.51	ดีมาก
2. ผลลัพธ์การค้นหาเป็นข้อมูลปัจจุบันและเป็นประโยชน์ต่อผู้ใช้	4.00	0.65	ดี
3. ผลลัพธ์การค้นหาตรงตามความต้องการของผู้ใช้	4.20	0.68	ดี

4. ระบบสามารถเข้าถึงเอกสารความรู้ด้านไม้ผลเศรษฐกิจ	3.93	0.59	ดี
โดยรวม	4.18	0.32	ดี

จากตารางที่ 4.21 พบว่า ผู้เชี่ยวชาญมีความคิดเห็นด้วยเกี่ยวกับความสามารถของระบบสืบค้นด้านการประเมินความสามารถของระบบอยู่ในระดับดีมาก 1 ข้อ คือ ความสามารถของส่วนนำเข้าข้อคำถามตรงตามความต้องการของท่านในระดับใด ($\overline{X}$ = 4.60) และอยู่ในระดับดี จำนวน 3 ข้อ โดยเรียงลำดับค่าเฉลี่ยจากมาก ไปหาน้อย ดังนี้ ผลลัพธ์การค้นหาตรงตามความต้องการของผู้ใช้ ($\overline{X}$ = 4.20) ผลลัพธ์การค้นหาเป็นข้อมูลปัจจุบันและเป็นประโยชน์ต่อผู้ใช้ ($\overline{X}$ = 4.00) และ ระบบสามารถเข้าถึงเอกสารความรู้ด้านไม้ผลเศรษฐกิจ ($\overline{X}$ = 3.93)

ตารางที่ 4.22 ความคิดเห็นเกี่ยวกับด้านการประเมินความถูกต้องของระบบ

ความสามารถของระบบสืบค้น ด้านการประเมินความถูกต้อง	$\overline{X}$	S.D.	ระดับความคิดเห็น
1. ความถูกต้องของข้อมูลสำหรับการค้นหาแบบหนึ่งคำค้น	4.33	0.72	ดี
2. ความถูกต้องของข้อมูลสำหรับการค้นหาแบบสองคำค้น	4.00	0.65	ดี
3. ความถูกต้องของข้อมูลสำหรับการค้นหาแบบสามคำค้น	4.00	0.76	ดี
4. ความถูกต้องของข้อมูลสำหรับการค้นหาแบบมากกว่าสามคำค้น	3.87	0.64	ดี
โดยรวม	4.05	0.61	ดี

จากตารางที่ 4.22 พบว่า ผู้เชี่ยวชาญมีความคิดเห็นด้วยเกี่ยวกับความสามารถของระบบสืบค้นด้านการประเมินความถูกต้องอยู่ในระดับดีทุกข้อ โดยเรียงลำดับค่าเฉลี่ยจากมากไปหาน้อย ดังนี้ ความถูกต้องของข้อมูลสำหรับการค้นหาแบบหนึ่งคำค้น ($\overline{X}$ = 4.33) ความถูกต้องของข้อมูลสำหรับการค้นหาแบบสองคำค้น ความถูกต้องของข้อมูลสำหรับการค้นหาแบบสามคำค้น ($\overline{X}$ = 4.00) และความถูกต้องของข้อมูลสำหรับการค้นหาแบบมากกว่าสามคำค้น ($\overline{X}$ = 3.87)

ตารางที่ 4.23 ความคิดเห็นเกี่ยวกับ ด้านการประเมินความง่ายในการใช้งาน

ความสามารถของระบบสืบค้น ด้านการประเมินความง่ายในการใช้งาน	$\overline{X}$	S.D.	ระดับความคิดเห็น

ตารางที่ 4.23 ความคิดเห็นเกี่ยวกับ ด้านการประเมินความง่ายในการใช้งาน

ความสามารถของระบบสืบค้น ด้านการประเมินความง่ายในการใช้งาน	$\overline{X}$	S.D.	ระดับความคิดเห็น
1. ขนาด สี รูปแบบตัวอักษรมีความเหมาะสม	4.53	0.52	ดีมาก
2. การออกแบบภาพกราฟิก มีความชัดเจน เหมาะสม	3.93	0.59	ดี
3. การจัดวางองค์ประกอบของระบบดูเหมาะสม	4.80	0.41	ดีมาก
4. การเชื่อมโยงไปยังเอกสารความรู้ที่ค้นพบมีความเหมาะสม	4.40	0.63	ดี
5. รูปแบบของการระบุคำค้นสะดวกต่อการใช้งาน	4.87	0.35	ดีมาก
6. ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้	4.07	0.59	ดี
โดยรวม	4.43	0.19	ดี

จากตารางที่ 4.23 พบว่า ผู้เชี่ยวชาญมีความคิดเห็นด้วยเกี่ยวกับความสามารถของระบบสืบค้น ด้านการประเมินความง่ายในการใช้งานอยู่ในระดับดีมาก จำนวน 3 ข้อ โดยเรียงลำดับค่าเฉลี่ยจากมากไปหาน้อย ดังนี้ รูปแบบของการระบุคำค้นสะดวกต่อการใช้งาน ($\overline{X}$ = 4.87) การจัดวางองค์ประกอบของระบบดูเหมาะสม ($\overline{X}$ = 4.80) ขนาด สี รูปแบบตัวอักษรมีความเหมาะสม ($\overline{X}$ = 4.53) และอยู่ในระดับดี 3 ข้อ โดยเรียงลำดับค่าเฉลี่ยจากมากไปหาน้อย ดังนี้ การเชื่อมโยงไปยังเอกสารความรู้ที่ค้นพบมีความเหมาะสม ($\overline{X}$ = 4.40) ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้ ($\overline{X}$ = 4.07) การออกแบบภาพกราฟิก มีความชัดเจน เหมาะสม ($\overline{X}$ = 3.93)

ตารางที่ 4.24 ความคิดเห็นเกี่ยวกับด้านการประเมินประสิทธิภาพของระบบ

ความสามารถของระบบสืบค้น ด้านการประเมินประสิทธิภาพของระบบ	$\overline{X}$	S.D.	ระดับความคิดเห็น
1. สามารถค้นหาข้อมูลได้อย่างรวดเร็ว	4.00	0.65	ดี
2. ระบบมีความเสถียรภาพในการเข้าใช้งาน	4.13	0.74	ดี
3. ระบบมีความแม่นยำในการค้นหา	3.80	0.77	ดี
โดยรวม	3.98	0.43	ดี

จากตารางที่ 4.24 พบว่า ผู้เชี่ยวชาญมีความคิดเห็นด้วยเกี่ยวกับความสามารถของระบบสืบค้นด้านการประเมินประสิทธิภาพของระบบอยู่ในระดับดีทุกข้อ โดยเรียงลำดับค่าเฉลี่ยจากมากไปหาน้อยดังนี้ ระบบมีความเสถียรภาพในการเข้าใช้งาน ($\overline{X}$ = 4.13) สามารถค้นหาข้อมูลได้อย่างรวดเร็ว ($\overline{X}$ = 4.00) และระบบมีความแม่นยำในการค้นหา ($\overline{X}$ = 3.80)

ตารางที่ 4.25 ความคิดเห็นเกี่ยวกับด้านการรักษาความปลอดภัย

ความสามารถของระบบสืบค้น ด้านการรักษาความปลอดภัย	$\overline{X}$	S.D.	ระดับ ความคิดเห็น
1. การกำหนดรหัสผู้ใช้ และรหัสผ่านในการตรวจสอบผู้ดูแลระบบ	3.27	0.70	ปานกลาง
2. การตรวจสอบสิทธิ์ก่อนการใช้งานของผู้ดูแลระบบ	3.53	0.64	ดี
3. การป้องกันการล้มเหลวของข้อมูลในระบบ	3.33	0.49	ปานกลาง
โดยรวม	3.38	0.49	ปานกลาง

จากตารางที่ 4.25 พบว่า ผู้เชี่ยวชาญมีความคิดเห็นด้วยเกี่ยวกับความสามารถของระบบสืบค้นด้านการรักษาความปลอดภัยอยู่ในระดับดี 1 ข้อ คือ การตรวจสอบสิทธิ์ก่อนการใช้งานของผู้ดูแลระบบ ($\overline{X}$ = 3.53) และอยู่ในระดับปานกลาง จำนวน 2 ข้อ ได้แก่ การป้องกันการล้มเหลวของข้อมูลในระบบ ($\overline{X}$ = 3.33) และการกำหนดรหัสผู้ใช้ และรหัสผ่านในการตรวจสอบผู้ดูแลระบบ ($\overline{X}$ = 3.27)

4.6 อภิปรายผลการวิจัย

ระบบที่พัฒนาขึ้นสามารถค้นหาเอกสารความรู้ สามารถเข้าถึงความรู้ทางการเกษตรได้ มีค่าความแม่นยำร้อยละ 96.84 ค่าการจำได้ร้อยละ 83.54 และอัตราการรู้จำร้อยละ 87.55 ซึ่งการประเมินความสามารถในการใช้งานระบบสืบค้นเชิงความหมายสำหรับเอกสารความรู้ด้านไม้ผลเศรษฐกิจของประเทศไทยโดยผู้เชี่ยวชาญ มีผลประเมินโดยรวมอยู่ในระดับดีทุกด้าน เป็นไปตามสมมุติฐานที่ตั้งไว้ เนื่องจากการพัฒนาระบบสืบค้นเชิงความหมาย ผู้วิจัยได้ทำการศึกษาหลักการ ทฤษฎี เอกสารงานวิจัยที่เกี่ยวข้อง ปัญหาที่เกี่ยวกับการเข้าถึงความรู้ทางการเกษตร และได้มีการดำเนินการวิจัยที่ใช้วิธีการพัฒนาระบบซึ่งประยุกต์ตามกระบวนการของวงจรการพัฒนาระบบ SDLC (System Development Life Cycle) ซึ่งการค้นหาจากระบบส่งผลให้การค้นหาเอกสารมีความถูกต้อง แม่นยำตรงกับความต้องการของผู้ใช้ ทั้งนี้ระบบสืบค้นเชิงความหมายที่นำเสนอจะมีประโยชน์อย่างยิ่ง หากเกษตรกรหรือผู้

ที่สนใจข้อมูลเอกสารความรู้การเกษตรโดยเฉพาะไม้ผลเศรษฐกิจ นำเอกสารความรู้ที่ค้นพบมาช่วยในการเพิ่มผลผลิต ลดต้นทุน หรือแก้ไขปัญหาจากการผลิตให้มีความสอดคล้องและตอบสนองความต้องการของการส่งออก นอกจากนี้ระบบยังช่วยให้เกษตรกรได้ความรู้จากเอกสารในการแก้ไขปัญหาได้ทันท่วงทีต่อสภาวะเหตุการณ์ไม้ผลเศรษฐกิจที่เกิดขึ้นในปัจจุบันโดยใช้ข้อมูลประกอบการตัดสินใจได้อย่างสะดวกและรวดเร็ว ซึ่งระบบที่นำเสนอจะเป็นฟันเฟืองอันหนึ่งที่จะขับเคลื่อนการเกษตรของไทยไปสู่การเกษตรแบบยั่งยืน ซึ่งการวิจัยครั้งนี้มีการนำเสนอเทคนิควิธีการสกัดข้อมูลโดยใช้สัมประสิทธิ์แจ็คการ์ดสำหรับวัดความคล้ายคลึงของคำค้น มีความสอดคล้องกับการวิจัยของจิรัสย์ มาลาวงษ์ และอานนท์ รุ่งสว่าง (2547) ได้วิจัยการเพิ่มประสิทธิภาพระบบสืบค้นเนมเพจด้วยการวิเคราะห์ชื่อที่ปรากฏเด่นชัดเป็นการเปรียบเทียบความใกล้เคียงระหว่างคำถามของผู้ใช้และตัวแทนชื่อเอกสารแต่ละอันในฐานข้อมูลทั้งหมด เอกสารที่มีค่าความใกล้เคียงกับคำถามของผู้ใช้มากที่สุดจะเป็นคำตอบของการสืบค้น ระบบสืบค้นข้อมูลทั่วไปนิยมใช้หลักการของความถี่ของคำที่ปรากฏในเอกสารรวมกับน้ำหนักของคำในเอกสารเป็นตัวแปรที่ระบุความใกล้เคียงของเอกสารและคำถามของผู้ใช้ แต่ระบบสืบค้นเนมเพจในงานวิจัยนี้ไม่สามารถใช้ตัวแปรที่ระบุข้างต้นได้เนื่องจากความถี่ของคำในเอกสารได้ถูกวิเคราะห์และแสดงเป็นชื่อเอกสารที่ปรากฏเด่นชัดแทน ทำให้ตัวแปรความถี่ของคำไม่สามารถนำมาระบุความใกล้เคียงได้ ในงานวิจัยนี้จึงใช้ฟังก์ชันในการเปรียบเทียบความใกล้เคียงแจ็คการ์ด (Jaccard Similarity Function) ฟังก์ชันดังกล่าวนิยมใช้ในการเปรียบเทียบความใกล้เคียงของกลุ่มข้อมูลในกระบวนการจัดกลุ่มข้อมูล (Data Clustering) (Guha et al, 1998) ฟังก์ชันนี้สามารถระบุค่าความใกล้เคียงของข้อมูล 2 ชุดได้อย่างมีประสิทธิภาพโดยไม่ต้องใช้ค่าตัวแปรความถี่ของข้อมูล จากผลการทดลองพบว่าการวิเคราะห์ชื่อเอกสารที่ปรากฏเด่นชัดสามารถค้นหาชื่อตัวแทนเอกสารได้อย่างถูกต้อง ส่งผลให้ความแม่นยำของระบบสูงขึ้นและยังใช้ฐานข้อมูลที่มีขนาดเล็กกว่าระบบสืบค้นเนมเพจทั่วไปอีกด้วย และสอดคล้องกับการวิจัยของ Niwattanakul (2008) ได้กล่าวไว้ว่า ขั้นตอนการสืบค้นคำ เริ่มต้นจากการนำเข้าคำสืบค้นจากผู้ใช้มาเทียบกับคำในดัชนีในคลังคำ ในกรณีคำที่เข้ามาตรงกับคำดัชนีในคลังคำ คำจะถูกนำมาใช้เป็นรายการคำหลักที่ใช้ในการค้นหาทันที แต่ในกรณีที่คำสืบค้นที่เข้ามาไม่ตรงกับคำดัชนีใด ๆ ในคลังคำ ก็จะทำกระบวนการวัดความคล้ายคลึงกันมากที่สุดของคำที่เข้ามากับคำในคลังคำซึ่งเก็บเป็นโครงสร้าง ได้แก่ คำหลัก คำเหมือน คำกว้างกว่า (Broader Term – BT) คำแคบกว่า (Narrower Term – NT) และ คำเกี่ยวข้อง (Related Term – RT) โดยใช้ Jaccard Similarity Function ค่าความคล้ายคลึงกันที่คำนวณต้องอยู่ในเกณฑ์ที่รับได้ ซึ่งมีค่าเท่ากับ 0.75 จะทำให้ได้คำหลักสำหรับการสืบค้น

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ

จากการศึกษาวิจัยการสืบค้นความรู้ไม้ผลเศรษฐกิจไทยโดยใช้เทคโนโลยีเว็บเชิงความหมาย ผลการวิจัยและข้อเสนอแนะมีดังนี้

5.1 สรุปผลการวิจัย

งานวิจัยนี้ได้ศึกษาและวิเคราะห์ขั้นตอนวิธีการค้นหาความรู้ไม้ผลเศรษฐกิจโดยใช้ออนโทโลยีอธิบายความหมาย ความสัมพันธ์และโครงสร้างของข้อมูลผ่านภาษาโอดับเบิ้ลยูแอล (Web Ontology Language : OWL) มีการใช้สัมประสิทธิ์แจ็คการ์ดสำหรับวัดความคล้ายคลึงของคำค้น ออกแบบขั้นตอนวิธีการสำหรับการค้นหาเชิงความหมาย และพัฒนาระบบสืบค้นเชิงความหมายสำหรับการเข้าถึงความรู้ไม้ผลเศรษฐกิจของประเทศไทย ได้ผลสรุปดังนี้

- **การออกแบบฐานความรู้ไม้ผลเศรษฐกิจไทยด้วยตัวแบบออนโทโลยี**

การออกแบบฐานความรู้ไม้ผลเศรษฐกิจด้วยตัวแบบออนโทโลยี มีโครงสร้างประกอบไปด้วย 8 คลาส (Class) 9 ออฟเจคพร็อพเพอร์ตี้ (Object properties) 34 ดาต้าไทป์พร็อพเพอร์ตี้ (Datatype properties) และคลังคำที่จัดการโดยใช้ตัวแบบเอสเคโอเอส (SKOS) คลาสหลักที่สร้างขึ้นมาจำนวน 8 คลาส ได้แก่ Economy_Fruit เป็นคลาสแทนข้อมูลไม้ผลเศรษฐกิจ Planting_and_Maintenance เป็นคลาสแทนข้อมูลการปลูกและบำรุงรักษา Diseases_and_Insects เป็นคลาสแทนข้อมูลโรคและแมลง Chemicals เป็นคลาสแทนข้อมูลสารเคมี Transformation เป็นคลาสแทนข้อมูลการแปรรูป Machines_and_Equipments เป็นคลาสแทนข้อมูลเครื่องจักรหรือเครื่องมือ Marketing_and_Exporting เป็นคลาสแทนข้อมูลการตลาดและการส่งออก และคลาส Documents ซึ่งเป็นคลาสแทนข้อมูลเอกสารความรู้ไม้ผลเศรษฐกิจ

สำหรับออฟเจคพร็อพเพอร์ตี้จำนวน 9 ออฟเจคพร๊อพเพอร์ตี้ (Object Properties) ได้แก่ hasDocument เป็นพร็อพเพอร์ตี้ที่เชื่อมโยงความสัมพันธ์ระหว่างคลาส skos:Concept และ Document, พร็อพเพอร์ตี้ hasTerms สำหรับเชื่อมโยงความสัมพันธ์ระหว่างคลาส Document และคลาส skos:Concept, พร็อพเพอร์ตี้ hasDiseases_Insects สำหรับเชื่อมโยงความสัมพันธ์ระหว่างคลาส Economy_Fruit และคลาสDiseases_and_Insects, พร็อพเพอร์ตี้ hasMarketing_Exporting สำหรับเชื่อมโยงความสัมพันธ์ระหว่างคลาส Economy_Fruit และ Marketing_and_Exporting, พร็อพเพอร์ตี้ hasPlanting_Maintenance สำหรับเชื่อมโยงคลาสความสัมพันธ์ระหว่างคลาส Economy_Fruit และคลาส Planting_and_Maintenance, พร็อพเพอร์ตี้ hasTransformation สำหรับเชื่อมโยงความสัมพันธ์ระหว่างคลาส Economy_Fruit และคลาส

Transformation, พร็อพเพอร์ตี้ useChemicals สำหรับเชื่อมโยงความสัมพันธ์ระหว่างคลาส Diseases_and_Insects และคลาส Chemicals, พร็อพเพอร์ตี้ useEquipments สำหรับเชื่อมโยงความสัมพันธ์ระหว่างคลาส Transformation และคลาส Machines_and_Equipments และพร็อพเพอร์ตี้ useMachines_Equipments สำหรับเชื่อมโยงความสัมพันธ์ระหว่างคลาส Planting_and_Maintenance และคลาส Machines_and_Equipments คือ โดยมีดาต้าไทป์พร็อพเพอร์ตี้ 34 ดาต้าไทส์พร็อพเพอร์ตี้ที่ครอบคลุมความรู้ไม้ผลเศรษฐกิจ

ในงานวิจัยนี้ได้พัฒนาคลังคำ ซึ่งประยุกต์โดยใช้ตัวแบบเอสเคโอเอส (Simple Knowledge Organization System: SKOS) ที่เป็นแหล่งรวบรวมคำศัพท์ต่าง ๆ โดยคำศัพท์แต่ละคำมีการจัดความสัมพันธ์ในลักษณะลำดับชั้น (Hierarchical Relation) ที่มีความหมายเหมือนกัน แต่เขียนต่างกัน โดยเน้นไปที่คำหลัก คำพ้อง คำกว้างกว่า (Broader Term-BT) คำแคบกว่า (Narrower Term-NT) และ คำเกี่ยวข้อง (Related Term-RT) มีคำศัพท์ที่เกี่ยวข้องกับไม้ผลเศรษฐกิจจำนวน 1,833 คำ

- **การวัดความคล้ายคลึงของคำค้น**

ขั้นตอนวิธีการวัดความคล้ายคลึงของคำค้นกับคำที่อยู่ในคลังคำโดยคำนวณสัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดในภาษาโปรล็อก และได้มีการทดสอบการวัดความคล้ายคลึงของคำปกติและคำที่มีข้อผิดพลาด ในการทดสอบใช้ค่าความแม่นยำ (Precision) ค่าการจำได้ (Recall) และค่าอัตราการรู้จำ (F-measure) ในการทดสอบกรณีพิมพ์ผิดหรือสะกดผิดพบว่าการวัดความคล้ายคลึงมีค่าความแม่นยำร้อยละ 83.33 ค่าการจำได้ร้อยละ 100 และค่าอัตราการรู้จำร้อยละ 90.90 การทดสอบกรณีพิมพ์ตกพบว่ามีค่าความแม่นยำร้อยละ 83.33 ค่าการจำได้ร้อยละ 100 และค่าอัตราการรู้จำร้อยละ 90.90 และการทดสอบกรณีพิมพ์เกินพบว่ามีค่าความแม่นยำร้อยละ 100 ค่าการจำได้ร้อยละ 100 และค่าอัตราการรู้จำร้อยละ 100 ดังนั้น วิธีการวัดความคล้ายคลึงโดยใช้สัมประสิทธิ์ความคล้ายคลึงของแจ็คการ์ดจะให้ผลดีในกรณีพิมพ์เกิน แต่ถ้าหากเป็นกรณีพิมพ์ผิด สะกดผิด หรือพิมพ์เกิน จะมีค่าความแม่นยำมากกว่าร้อยละ 80

- **การสืบค้นเชิงความหมาย**

การสืบค้นเชิงความหมายจะมีการให้ความสำคัญกับการวัดความคล้ายคลึงของคำค้น และมีการคำนวณระยะทางที่สั้นที่สุดจากคำที่อยู่ในโครงสร้างที่เก็บในคลังคำ ประกอบด้วย 3 ขั้นตอน ได้แก่ 1) ขั้นตอนของการกำหนดความสำคัญของคำค้น โดยจะกำหนดให้คำที่ไม่มีคำกว้างกว่ามีค่าความลึกเป็น 1 และ กำหนดให้คำที่เหมือนกันหรือเกี่ยวข้องมีค่าเท่ากัน จากนั้นกำหนดค่าคาดหวังให้คำแคบกว่ามีค่าเพิ่มหนึ่ง และคำนวณค่าความสำคัญจากผลรวมค่าคาดหวังหารด้วยจำนวนความคาดหวัง 2) คำนวณหาค่าน้ำหนักของคำค้น โดยในขั้นตอนนี้จะทำการปรับค่าน้ำหนักของคำค้นให้สามารถนำมาวิเคราะห์ความสำคัญได้ ซึ่งการปรับค่าความสำคัญให้ได้ผลรวมของค่าน้ำหนักเท่ากับหนึ่ง 3) หาเส้นทางสั้นที่สุดตามโครงสร้าง โดยกำหนดให้คำที่เหมือนกันหรือเกี่ยวข้องมีค่าระยะทางเป็น 0 กำหนด

คำที่แคบกว่าหรือกว้างกว่ามีค่าระยะทางเป็น 1 และคำนวณค่าระยะทางโดยการหาผลรวมในเส้นทางที่สั้นที่สุดจากจุดหนึ่งไปยังอีกจุด หลังจากที่ได้กลุ่มคำค้นและรายการเอกสารความรู้ที่เกี่ยวข้องมาแล้ว จะทำการเปรียบเทียบความคล้ายคลึงของคำค้นกับเอกสารแต่ละรายการ และเรียงลำดับค่าความคล้ายคลึงจากมากไปหาน้อยสำหรับแสดงผลลัพธ์

- **ผลการประเมินระบบสืบค้นเชิงความหมาย**

การทดสอบสืบค้นเชิงความหมายมีการทดสอบผ่านส่วนติดต่อกับผู้ใช้งานบนเว็บบราวเซอร์จากฐานความรู้ที่สร้างขึ้น ซึ่งสามารถเข้าไปสืบค้นข้อมูลได้ที่ http://www.thaifo.net มีผลการประเมิน ดังนี้

- *การประเมินความถูกต้องของระบบสืบค้น*

ผลการทดสอบประสิทธิภาพการค้นคืนโดยการเปรียบเทียบระหว่างการสืบค้นแบบจับคู่กับการสืบค้นเชิงความหมายพบว่าการสืบค้นเชิงความหมายจะให้ผลการสืบค้นที่ดีกว่า โดยจะเห็นได้จากผลการทดสอบการสืบค้นเชิงความหมายมีค่าความแม่นยำร้อยละ 96.84 ค่าการจำได้ร้อยละ 83.54 และค่าอัตราการรู้จำร้อยละ 87.55 ในขณะที่ผลการทดสอบการสืบค้นแบบจับคู่มีค่าความแม่นยำร้อยละ100.00 ค่าการจำได้ร้อยละ 55.77 และค่าอัตราการรู้จำร้อยละ 67.31 ซึ่งสรุปได้ว่าการสืบค้นแบบจับคู่จะทำให้ได้ผลที่ตรงกับความต้องการมาก แต่จะมีข้อเสียที่ไม่สามารถค้นคืนเอกสารอื่น ๆ ที่มีความเกี่ยวข้องได้ดังจะเห็นการสืบค้นแบบจับคู่จะมีค่าการจำได้เพียงร้อยละ 55.77

- *การประเมินความสามารถของระบบสืบค้น*

การประเมินความสามารถของระบบสืบค้นเชิงความหมายมีการประเมินใน 5 ด้าน ประกอบด้วย 1) ด้านการประเมินความสามารถของระบบ (Functional requirement test) 2) ด้านการประเมินความถูกต้องของระบบ (Functional test) 3) ด้านการประเมินความง่ายในการใช้งาน (Usability test) 4) ด้านการประเมินประสิทธิภาพของระบบ (Performance test) 5) ด้านการประเมินด้านการรักษาความปลอดภัย (Security test) พบว่า ผู้ตอบแบบสอบถามประเมินความสามารถระบบสืบค้นส่วนใหญ่เป็นเพศชาย ระดับการศึกษาปริญญาโท มีความคิดเห็นเกี่ยวกับความสามารถของระบบสืบค้นเชิงความหมายโดยรวม อยู่ในระดับดี 4 ด้าน ได้แก่ ด้านการประเมินความง่ายในการใช้งาน ด้านการประเมินความสามารถของระบบ ด้านการประเมินความถูกต้องของระบบ ด้านการประเมินประสิทธิภาพของระบบ และอยู่ในระดับปานกลาง 1 ด้าน ได้แก่ ด้านการประเมินด้านการรักษาความปลอดภัย

5.2 ข้อเสนอแนะ

สำหรับเทคนิควิธีการสกัดข้อมูลการวัดความคล้ายด้วยสัมประสิทธิ์แจ๊คการ์ด และขั้นตอนวิธีค้นหาเชิงความหมายนี้ถูกออกแบบมาเพื่อพัฒนาระบบสืบค้นเชิงความหมายในงานวิจัยนี้เท่านั้น แต่ถ้าหากต้องการที่จะนำเอากระบวนวิธีนี้ไปประยุกต์เพื่อใช้ในการสืบค้นเชิงความหมายในด้านอื่น อาจต้องทำการศึกษาเพิ่มเติม และงานวิจัยในอนาคตควรจะมีการออกแบบและพัฒนาเทคนิควิธีการวัดความคล้ายคลึงกัน โดยทำการศึกษาวิธีการวัดความคล้ายเพื่อสรรหาวิธีการที่ดีที่สุดสำหรับการวัดความคล้ายของคำค้น ซึ่งจะได้คำค้นที่ถูกต้องถึงแม้ว่าผู้ใช้จะพิมพ์ผิดและหาขั้นตอนวิธีการค้นหาเชิงความหมายสำหรับนำไปพัฒนาระบบสืบค้นเชิงความหมายที่มีความถูกต้อง แม่นยำตรงกับความต้องการของผู้ใช้มากยิ่งขึ้น

บรรณานุกรม

กรมวิชาการ. (2545). เอกสารประกอบหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 การวิจัยเพื่อพัฒนาการเรียนรู้ตามหลักสูตรการศึกษาขั้นพื้นฐาน. กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว.

กานดา ศรอินทร์. (2548). ระบบสืบค้นข้อมูลเฉพาะเจาะจงโดยใช้คำเฉพาะและออนโทโลยี. การศึกษาอิสระ. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

กิสณะ ตันเจริญ. (2552). ไม้ผลยืนต้น และยางพารา. กรุงเทพฯ : กลุ่มส่งเสริมการผลิตไม้ผล กรมส่งเสริมการเกษตร กระทรวงเกษตรและสหกรณ์.

จริยา วิสิทธิ์พานิช และชาตรี สิทธิกุล. (2548). โรคและแมลงศัตรูสำคัญของลำไย ในลำไยคุณภาพ. วันแม่โจ้ : ศาสตร์แห่งลำไย 25-27 พฤศจิกายน 2548. เชียงใหม่ : มหาวิทยาลัยแม่โจ้. หน้า 11-14

จารุณี แซ่หลี. (2549). การค้นหาคำสำคัญในฐานข้อมูล โดยใช้การค้นข้อมูลฐานข้อมูล. วิทยานิพนธ์. สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

จิรัสย์ มาลาวงษ์ และอานนท์ รุ่งสว่าง. (2547). การเพิ่มประสิทธิภาพระบบสืบค้นเนมเพจด้วยการวิเคราะห์ชื่อที่ปรากฏเด่นชัด. วิทยานิพนธ์ วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเกษตรศาสตร์.

ชมพูนุช ภักติยานุวรรตน์. (2550). การบูรณาการออนโทโลยีอย่างมีความหมายจากแหล่งข้อมูลที่หลากหลาย. การศึกษาอิสระ. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

เดโชดม ภัทรศัย. (2543). 36 ผลไม้สมุนไพรไทย. กรุงเทพฯ : โปร-เอสเอ็มอี.

ธีรนุช เจริญกิจ และพาวิน มะโนชัย. (2548). การเก็บเก็บเกี่ยวและการจัดการหลังการเก็บเกี่ยว. ในลำไยคุณภาพ. วันแม่โจ้ : ศาสตร์แห่งลำไย 25-27 พฤศจิกายน 2548. เชียงใหม่ : มหาวิทยาลัยแม่โจ้. หน้า 47-50.

นคร โคตรโสภา .(2549). วิธีการสืบค้นเอกสารเชิงความหมายบนข้อมูล XML. การศึกษาอิสระ. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

นาถธิชา เกษรพันธ์. (2554). การพัฒนาออนโทโลยีสำหรับระบบให้คำแนะนำอาจารย์ ผู้เชี่ยวชาญตามหัวข้องานวิจัยของนักศึกษา. การค้นคว้าแบบอิสระ วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมซอฟต์แวร์ มหาวิทยาลัยเชียงใหม่.

บุญชม ศรีสะอาด. (2545) การวิจัยเบื้องต้น. พิมพ์ครั้งที่ 7. กรุงเทพฯ : สุวีริยาสาส์น.

บุญรัตน์ เผดิมรอด. (2551). วิธีการใหม่สำหรับสืบค้นแบบหลายความสัมพันธ์บนองค์ความรู้ในฐานข้อมูลเชิงสัมพันธ์. วิทยานิพนธ์. วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเกษตรศาสตร์.

ปฏิคม ทองจริง. (2552). ออนโทโลยีสำหรับการรวมข้อมูลเชิงความหมายของความรู้ด้านสมุนไพรไทย. วิทยานิพนธ์. วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยธรรมศาสตร์

ประสิทธิ์ ประทุมรัตน์. (2550). การจัดเก็บและสืบค้นรูปภาพเชิงความหมายโดยใช้ออนโทโลยี. การศึกษาอิสระ. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

ปิยวัฒน์ ทองแก้ว และสมชาย ปราการเจริญ. (2553). การพัฒนาระบบจัดเก็บและค้นคืนสารนิพนธ์ทางเทคโนโลยีสารสนเทศ โดยใช้คำศัพท์ควบคุมที่มีโครงสร้างแบบธิซอรัสบนแนวคิดออนโทโลยีสำหรับห้องสมุดดิจัล. เอกสารประกอบการประชุมวิชาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 6 (The 6th National Conference on Computing and Information Technology: NCCIT), หน้า 784-789.

พงษ์ศักดิ์ อังกสิทธิ์ และคณะ. (2542). ลำไย : ไม้ผลเศรษฐกิจสำคัญเพื่อพัฒนาอุตสาหกรรม. พิมพ์ครั้งที่ 1. เชียงใหม่ : คณะเกษตรศาสตร์, มหาวิทยาลัยเชียงใหม่. 137 หน้า.

พนิดา ตันศิริ. (2553). เว็บเชิงความหมายของเว็บ 3.0. วารสารนักบริหาร, หน้า 48-55.

พาวิน มะโนชัย, ยุทธนา เขาสุเมรุ, ชิติ ศรีตนทิพย์ และสันติ ช่างเจรจา. (2547). เทคโนโลยีการผลิตลำไย. พิมพ์ครั้งที่ 1 . กรุงเทพฯ : ฟิสิกส์เซ็นต์เซ็นเตอร์. 128 หน้า

เพ็ญพรรณ อัศวนพเกียรติ. (2547). ระบบสำหรับสืบค้นและรวบรวมข้อมูลทางชีววิทยาโดยใช้ชีวออนโทโลยี. วิทยานิพนธ์. วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเกษตรศาสตร์.

ภรศิษฐ์ กุหลาบทิพย์. (2550). การค้นหาข้อมูลบนเว็บเชิงความหมายโดยใช้ฐานความรู้ WORDNET. การศึกษาอิสระ. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

มนัสนันท์ ปัญญามี. (2552). การจัดกลุ่มผลการสืบค้นเอกสารบนเว็บเพื่อการสืบค้นเชิงความหมาย. วิทยานิพนธ์. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

มารุต บูรณรัช และ เทพชัย ทรัพย์นิธิ. (2550). Semantic Web. ออนไลน์: http://text.hlt.nectec.or.th/ ontology/content/what-is-semantic-km

ยุทธนา สุดเจริญ. (2548). การแยก และพิสูจน์เอกลักษณ์ของสารต้านอนุมูลอิสระกลุ่มฟีนอลลิกจากของเหลือทิ้งจากผลไม้ไทย. โครงการปริญญาเอก เวชศาสตร์เขตร้อน, กรุงเทพฯ : มหาวิทยาลัยมหิดล.

โรสริน อัคนิจ, ธนิต พุทธพงศ์ศิริพร, น้ำผน ลำดับวงศ์ และอารีย์ ธัญกิจจานุกิจ. (2554). การพัฒนาออนโทโลยี เพื่อการจัดการความรู้ด้านการแปรรูปข้าว. วารสารเกษตร, 27(3) : 267-274.

วันดี กฤษณพันธ์. (2541). สมุนไพรน่ารู้ โครงการสมุนไพรเพื่อการพึ่งตนเอง. พิมพ์ครั้งที่ 3. กรุงเทพฯ : ประพันธ์สาส์น.

วันทนา บัวทรัพย์ และ มนตรี วงศ์รักษ์พานิช. (2548). เอกสารประกอบการฝึกอบรมหลักสูตรเทคโนโลยี : การปลูกทุเรียน. กรุงเทพฯ : กรมส่งเสริมการเกษตร กระทรวงเกษตรและสหกรณ์.

วิลาศ วูวงศ์. (2548). ว่าด้วยเมทาดาทา. ออนไลน์: http://www.limahidol.ac.th/seminar22nd/present/vilas.pdf

วีรนันท์ มาลาศิลป์. (2550). การบูรณาการข้อมูลอย่างมีความหมายจากแหล่งข้อมูลที่หลากหลายโดยคำอธิบายข้อมูลออนโทโลยี. การศึกษาอิสระ. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

ศุภชัย ตั้งวงศ์ศานต์. (2553). ระบบการจัดเก็บและการสืบค้นสารสนเทศด้วยคอมพิวเตอร์. กรุงเทพฯ : พิทักษ์การพิมพ์.

ศุภลักษณ์ กลับน่วม และอัญชลี พัดมีเทศ. (2547). แผ่นพับเผยแพร่ที่ 172 : โรคทุเรียน. กรุงเทพฯ : กรมส่งเสริมการเกษตร กระทรวงเกษตรและสหกรณ์.

สมมณี ลูซะวงษ์. (2553). การเข้าถึงฐานข้อมูลบนพื้นฐานของออนโทโลยีด้วยวิธีการปรับเปลี่ยนคำสั่งการสืบค้น. วิทยานิพนธ์. วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น.

สราวุธ อ้อยศรีสกุล. (2544). เริ่มคิด เริ่มสร้าง เริ่มใช้ XML. กรุงเทพฯ : วิตตี้ กรุ๊ปจำกัด.

สำนักงานเศรษฐกิจการเกษตร. (2554). สถิติการเกษตรของประเทศไทย ปี 2553. กระทรวงเกษตรและสหกรณ์. ออนไลน์ : http://www.oae.go.th/download/download_journal/yearbook53.pdf

สุรชาติ เพชรแก้ว. (2542). ศักยภาพที่ดินสำหรับการปลูกมังคุดในพื้นที่ภาคใต้ของประเทศไทย สมบัติบางประการของดินปลูกมังคุดในภาคใต้ของประเทศไทย สมบัติบางประการของดินปลูกมังคุดในอำเภอนาหม่อม จังหวัดสงขลา. ภาควิชาธรณีศาสตร์ มหาวิทยาลัยสงขลานครินทร์.

อติพงศ์ จุลละโพธิ. (2548). เสิร์ชเอนจิ้น การค้นไอเดียบนเว็บที่ยังไม่ดีพอ. Retrieved Jan 18, 2013, Available URL : http://content.ee43.com/content/150

อรวรรณ อุไรเรืองพันธ์ และสมจิตร อาจอินทร์. (2552). การสรุปเอกสารเชิงความหมายโดยใช้ออนโทโลยี. เอกสารประกอบการประชุมวิชาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 5 (The 5th National Conference on Computing and Information Technology: NCCIT), หน้า 294-299.

อารีย์ ธัญกิจจานุกิจ. (2556). เครือข่ายสารสนเทศเกษตรไทย. ออนไลน์ : http://www.slideshare.net/ boonlert/thai-arguriculture

Antoniou, G and Van Harmelen, F. (2008). A Semantic Web primer. 2nd ed. London : MIT Press.

Berners-Lee, T. (2006). About the World Wide Web Consortium (W3C). Available: http://www.w3.org/Consortium

Berners-Lee, T., James, H., and Ora, L. (2001). The Semantic Web. Available: http://www.si.umich.edu/ ~rfrost/courses/si110/readings/in_out_and_beyond/semantic_web.pdf.

Brendan, E., En C., Chimezie, T.O., and Ozsoyoglu, Z.M. (2009). A Complete Translation from SPARQL into Efficient SQL. New York : Bill press.

Broughton, V. (2006). The need for a faceted classification as the basis of all methods of information retrieval. Aslib Proceedings: New Information Perspective, 58(1/2), 49-72.

Christopher, C. V. and Garrison, W, C. (1998). Predicting the performance of linearly Combined IR systems. In Proceedings of 21th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, New York, pp. 190-196.

Craswell, N., Hawking, D. and Robertson, S. (2001). Effective site finding using link anchor information. In Proceedings of 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, New York, pp. 250-257.

Duineveld, A.J., Stoter, R., Weiden, M. R., Kenepa, B. and Benjamins, V. R. (2000). Wondertools? A comparative study of ontological engineering tools. In the International Journal of Human-Computer Studies, July 2000 (in press).

FAO. (2013). AGROVOC, Food and Agricultural Organization of the United Nations. Retrieved Dec 20, 2012, Available URL: http://aims.fao.org/website/AGROVOC-Thesaurus/sub

Fensel, D. (2004). Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce. IEEE Intelligent Systems, 16(1), 8-14.

Fensel, D., Harmelen, F.V., Horrocks, I, McGuinness, D.L, and Patel-Schneider, P.F. (2001). OIL : An Ontology Infrastructure For the Semantic web. IEEE Intelligent Systems 2001. 2(16)

Fulvio, C. , Laura, F. (2005). SKOS -Simple Knowledge Organization System. Available: http://elite.polito.it

Gruber, T. R. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Online. Available: http://tomgruber.org/writing/onto-design.thml.

Guarino, G. (1998). “Formal Ontology in Information System.” 1st International Conference On Formal Ontology in Information System (FOIS’98). Trento, Italy. IOS Press, Amsterdam.

Guha, S., Rastogi, R., Shim, K. (1998). An efficient clustering algorithm for large databases. In Proc. of SIGMOD international Conf. on Management of data. New York, USA. pp. 73-84.

Hahn, U. and Mani, I. (2000). “The Challenges of Automatic Summarization,” IEE Computer Society Press, vol. 33, pp. 29-36.

Hart L., Emery P., Colomb B., Raymond K. and Taraporewalla S. (2004). OWL Full and UML 2.0 Compared. Available: http://www.Itee.uq.edu.au/~colomb/Papers/ UML-OWLLont04.03. 01.pdf

Hodge, G. (2000). Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files. Available: http://www.clir.orgpubs/reports/ pub91/contents.html

Hunter, D., Cagle, C., Gibbon, D., Ozu, N., Pinnock, J. and Spencer, P. (2002). คัมภีร์การใช้ XML ฉบับสมบูรณ์. (สุวัฒนา สุขสมจินต์, ผู้แปล). กรุงเทพฯ : SE-ED Education Public Company.

Hunter, J. (2001). MetaNet – A metadata term theasuarus to enable semantic Interoperability between metadata domains. Journal of Digital Information, 1(8), 42, Retrieved November 26, 2011, Available : http://jodi.tamu.edu/Articles/ v01/i08/Hunter/

Jacob, E. K. (2003). Ontologies and the Semantic Web. Special Section. Bulletin of the American Society for Information Science and Technology. pp. 19-22.

Klyne, G. and Carroll J. (2004). Resource Description Framwork (RDF) : Concepts and Abstract Syntax, W3C Candidate Recommendation 2004-02-10. Retrieved Nov 12, 2005, Available URL: http://www.w3.org/TR/rdf-concepts

McConnell, J. (2001). Analysis of Algorithms : An Active Learning Approach. Canada : Jones and Bartlett, pp. 163-168.

Miao, D., Duan, Q., Zhang, H. and Jiao, N. (2009). Rough set based hybrid algorithm for text classification. Expert Systems with Applications. 36(5): 9168-9174.

Nejdl, W., Wolpers, M. and Capelle, C. (2000). The RDF Schema Specification Revised. Paper presented at the Modellierung 2000, Germany.

Niwattanakul, S. (2008). Access to Knowledge Based-on an Ontology Model, Ph.D. thesis, University of La Rochelle, France.

Niwattanakul, S., Singthongchai, J., Naenudorn, E. and Wanapu, S. (2013). Using of Jaccard Coefficient for Keywords Similarity, In Proceedings of The International MultiConference of Engineers and Computer Scientists (IMECS), Hong Kong, pp. 380-384.

Oldakowski, R. (2004). RDQL Tutorial : This tutorial is part of the RAP-RDF API for PHP documentation. Retrieved Nov 20, 2006, Available URL: http:/sites.wiwiss.fu-berlin. De/suhl/bizer/rdfapi/tutorial/rdql_tutorial.htm

Oufaida H. and Omar N. (2009). Exploiting Semantic Web Technologies for Recommender Systems A Multi View Recommendation Engin. DTISI Laboratory, CERIST Research Center 03, Algeria.

Prud'hommeaux, E. and Seaborne, A. (2007). SPARQL Query Language for RDF. W3C Recommendation. Available : http://www.w3.org/TR2007/PR-rdf-sparql-query-20071112/

Seaborne, A. (2004). A Programmer's Introduction to RDQL. Retrieved Dec 12, 2012, Available URL: http://www.hpl.hp.com/semweb/doc/tutorial/RDQL

NECTEC. (2009). Semantic Web. Available: http://www.nectec.or.th/images/pdf/techtrends/62/itdig.pdf.

Smith, M. K., Welty, C. and McGuinness, D. L. (2004). OWL web ontology language guide. W3C Recommendation 20040210, W3C. Available: http://www.w3.org/TR/owl-guide/.

Studer, R., Benjamins V.R. and Fensel, D. (1998). Knowledge Engineering: Principles and Methods. Data and Knowledge Engineering. vol.25, pp. 161-197.

Uschold, M. and Gruninger, M. (1996). Ontologies: principles, methods and application. Knowledge Engineering Review, 11(2), 93-155.

Vong, H. L. (2003). Msc in Machine Learning and Data Ming Project. (Reinforcement Learning Directed Search Agent on Semantic Web). London : University of Bristol.

W3C. (2004). Resource description framework (RDF). Available: http://www.w3.org/RDF

W3C. (2006). W3C Semantic Web Activity. Available: http://www.w3org/2001/sw

W3C. (2009). Ontology. Available: http://en.wikipedia.org/wiki/World_Wide_Web_ Consortium. html

Wang, J., Miao, Z., Zhang, Y., and Zhou, B. (2009). Querying Heterogeneous Relational Database using SPARQL. USA : IEEE/ACIS.

ภาคผนวก

ภาคผนวก ก
ข้อมูลไม้ผลเศรษฐกิจไทย

ข้อมูลไม้ผลเศรษฐกิจของประเทศไทย

(1) ลำไย

- **ลักษณะทางพฤกษศาสตร์ (พงษ์ศักดิ์ อังกสิทธิ์ และคณะ, 2542)**

ชื่อภาษาไทย ลำไย

ชื่อภาษาอังกฤษ Longan

ชื่อวิทยาศาสตร์ Dimocarpus longen Lour. Euphoria longana Lamk. Nephelium longanum Camb.

วงศ์ Sapindaceae

ลำไยเป็นไม้ผลตระกูล Sapindaceae จำแนกได้เป็น 2 สายพันธุ์ (Species) ขึ้นอยู่กับลักษณะของลำต้น ผล และเมล็ด

ลำต้น มีขนาดลำต้นสูงปานกลางจนถึงขนาดใหญ่ ต้นที่ขยายพันธุ์ด้วยเมล็ดจะมีลำต้นตรงเมื่อเจริญเติบโตเต็มที่มีความสูงประมาณ 12.15 เมตร และถ้าหากเป็นต้นที่ขยายพันธุ์ด้วยการตอนกิ่งจะแตกกิ่งก้านสาขาใกล้ๆกับพื้น และถ้าได้รับการตัดแต่งกิ่งในขณะที่ต้นยังเล็กมักแตกลำต้นเทียมหลายต้น ลำต้นที่เกิดขึ้นไม่ค่อยเหยียดตรงมักเอนหรือโค้งงอเปลือกลำต้น ขรุขระมีสีเทาหรือสีเทาปนน้ำตาลแดงเป็นสะเก็ด

ใบ เป็นใบรวมที่ประกอบด้วยใบย่อยอยู่บนก้านใบร่วมกัน(Pinnately Compound Leaves) มีปลายใบเป็นคู่มีใบย่อย 3.-5 คู่ ความยาวใบ 20-30 เซนติเมตร ใบย่อยเรียงตัวสลับหรือเกือบตรงข้าม ความกว้างของใบย่อย 3-6 เซนติเมตร ยาว 7-15 เซนติเมตร รูปร่างใบเป็นรูปรีหรือรูปหอก ส่วนปลายใบและฐานใบค่อนข้างป้าน ใบด้านใบมีสีเขียวเข้มกว่า ด้านล่างสากเล็กน้อย ขอบใบเรียบไม่มีหยัก ใบเป็นคลื่นเล็กน้อย และเห็นเส้นแขนง (Vein) แตกออกมาจากเส้นกลางใบชัดเจนและมีจานวนมาก

ช่อดอก ส่วนมากเกิดจากตาที่ปลายยอด(Terminal Bud) บางครั้งอาจเกิดจากตาข้างของกิ่ง ช่อดอกยาวประมาณ 15-60 เซนติเมตร ช่อดอกขนาดกลางจะมีดอกย่อยประมาณ 3,000 ดอก ลักษณะช่อดอกลำไยแสดงดังภาพที่ ก.1 (พงษ์ศักดิ์และคณะ, 2542)

ภาพที่ ก.1 ลักษณะช่อดอกลำไย

ดอก มีสีขาวหรือขาวอมเหลืองมีขนาดเล็กเส้นผ่าศูนย์กลางประมาณ 6-8 มิลลิเมตร มีกลิ่นหอม ช่อดอกหนึ่ง ๆอาจมีดอก 3ชนิด (Polygamo-Monoecious) ดอกตัวผู้ (Staminate) ดอกตัวเมีย (Pistillate Flower) และดอกสมบูรณ์เพศ (Perfect Flower) ลักษณะที่คล้ายคลึงกันของดอกทั้ง 3 ชนิด คือกลีบดอกบาง 5 กลีบ สีขาว กลีบเลี้ยงหนาแข็ง 5 กลีบมีสีเขียวปนน้ำตาล

ผล มีผลทรงกลมเบี้ยว เปลือกสีน้ำตาลปนเหลืองหรือปนเขียว ผลสุกมีเปลือกสีเหลืองหรือสีน้ำตาลอมแดงผิวเปลือกเรียบหรือเกือบเรียบ มีตุ่มแบนๆปกคลุมที่ผิวเปลือกด้านนอกเนื้อลำไยเป็นเนื้อเยื่อพาเรนไคมาที่เจริญล้อมรอบเมล็ด(Outer Integument) และอยู่ระหว่างเปลือกกับเมล็ด ซึ่งมีสีขาวคล้ายวุ้น มีสีขาวขุ่น ใสหรือสีชมพูเรื่อๆ มีกลิ่นหอม รสหวาน แตกต่างกันไปตามพันธุ์

เมล็ด มีลักษณะกลมจนถึงแบน เมื่อยังไม่แก่มีสีขาวแล้วค่อยๆเปลี่ยนเป็นสีดำมันส่วนของเมล็ดที่ติดกับขั้วผล (Dragon’ eye) นี้จะมีขนาดเล็กหรือใหญ่ต่างกันไปตามพันธุ์ เมื่อผลแก่จัดถ้ายังไม่เก็บเกี่ยว placenta จะใหญ่ขึ้นเนื่องจาก Placenta ดูดอาหารขึ้นไปเลี้ยงเมล็ด ทำให้เนื้อเยื่อมีรสชาติจืดลง

- **ประเภทหรือกลุ่มของลำไย แบ่งออกเป็น 3 กลุ่มดังนี้**

1) ลำไยป่า ลักษณะต้นใหญ่ ขนาดผลเล็กมาก ขั้วผลบาง ใช้ในการขยายพันธุ์

2) ลำไยพื้นเมือง ลักษณะลำต้นใหญ่ ผลผลิตสูง ขนาดผลเล็ก ขั้วผลบาง ความหวานน้อย (ประมาณ 13.75 %) คุณภาพต่ำ เมล็ดใหญ่ ใช้เป็นต้นตอ

3) ลำไยปลูก ลักษณะลำต้นขนาดกลาง หรือขนาดใหญ่ ขั้วผลหนา มีความหวานสูง (ประมาณ 19.00-23.50 %) คุณภาพสูง เมล็ดเล็ก รับประทานสดหรือแปรรูป

ลำไยที่มีการปลูกในประเทศไทย สามารถแบ่งออกเป็น 2 กลุ่ม คือลำไยเครือและลำไยต้น (พาวิน มะโนชัย และคณะ, 2547)

1) ลำไยเครือหรือลำไยเถา (*Euphoria scandens* Winit. Kerr) เป็นลำไยที่มีลำต้นเลื้อยคล้ายเถาวัลย์ ทรงพุ่มคล้ายต้นเฟื่องฟ้า ลำต้นไม่มีแก่น ใบขนาดเล็กและสั้น ผลเล็ก ผิวสีชมพูปนน้ำตาล เมล็ดโต เนื้อผลบางมีกลิ่นคล้ายกำมะถัน ปลูกไว้สำหรับเป็นไม้ประดับมากกว่า พบได้ทั่วไปในแถบภาคตะวันออกของประเทศไทยแถวจังหวัดชลบุรี

2) ลำไยต้น สามารถแบ่งออกได้เป็น 3 ชนิด

2.1) ลำไยดั้งเดิม (Indigenous longan) พบได้ทั่วไปในป่า ปลูกจากเมล็ด เนื้อบางมาก แต่มีความสำคัญต่อการปรับปรุงพันธุ์

2.2) ลำไยพื้นเมือง (Common or Native longan) หรือลำไยกระดูก พบได้ในภาคเหนือและภาคตะวันออกเฉียงเหนือในแถบจังหวัดเชียงใหม่ จังหวัดลำพูน จังหวัดหนองคาย และจังหวัดอุบลราชธานี เป็นต้น เปลือกลำต้นและกิ่ง ขรุขระมาก ลำต้นตั้งตรงสูงประมาณ 20-30 เมตร ให้ผลผลิตสูง แต่ผลมีขนาดเล็กรูปร่างค่อนข้างกลม สีน้ำตาล เนื้อบางสีขาวใส คุณภาพผลค่อนข้างต่ำ ปัจจุบันใช้เป็นต้นตอสำหรับพันธุ์ปลูกทั่วไป

2.3) ลำไยกะโหลก เป็นลำไยพันธุ์การค้าที่นิยมปลูกกันทั่วไป ผลมีขนาดใหญ่ เนื้อหนา มีรสหวาน เมล็ดมีขนาดเล็ก มีหลายพันธุ์ดังต่อไปนี้ ชมพู ดอ ดอก้านแข็ง ดอก้านอ่อน ดอคำลาง ดอแจ้ คอดอยไชย ดอทาน้อย ดอน่าน ดอน้าผึ้ง ดอบ้านโฮ่ง 60 ดอใบดำดอใบหด ดอยอดขาว ดอยอดแดง ดอลุ่มน้ำปิง ดอสร้อย ดอสุขุม ดอหนองช้างคืน ดอหนานขาว ดอหลวง ดอหอม ดอ 13 แดงกลม นราภิรมย์ เบี้ยวเขียว เบี้ยวเขียวเชียงใหม่ เบี้ยวเขียวป่าเส้า ใบดา ใบหด ใบหยก เพชรเวียงพิงค์ เพชรสาคร เพชรสาครทะวาย แห้ง แห้วแคระ แห้วยอดขาว แห้วยอดแดง เหลือง อีดอ อีดอยอดเขียว อีดอยอดเขียวก้านอ่อน อีดอกยอดแดง อีแดง อีเบี้ยว อีเหลือง อีแห้ว

นอกจากนี้ ลำไยยังสามารถแบ่งออกได้อีก 3 ลักษณะตามระยะเวลาในการติดดอกออกผล คือ (พงษ์ศักดิ์และคณะ, 2542)

1. พันธุ์เบา ได้แก่ พันธุ์ดอ
2. พันธุ์ปานกลาง ได้แก่ พันธุ์แห้ว พันธุ์สีชมพู
3. พันธุ์หนัก ได้แก่ พันธุ์อีแดง เบี้ยวเขียว และพันธุ์พวงทอง

- **การปลูกลำไย**

1) การเลือกพื้นที่ ลำไยเป็นพืชที่เจริญเติบโตในดินแทบทุกชนิด แม้กระทั้งดินลูกรัง แต่ดินปลูกที่ให้ลำไยมีการเจริญเติบโตได้ดี คือดินร่วนปนทรายและดินตะกอน ซึ่งเกิดจากตะกอนดินกรวด หิน ดิน

ทราย อินทรีวัตถุที่น้ำพัดมาเกิดการทับถมของอินทรียวัตถุ ดินปลูกลำไยควรมีค่าความเป็นกรดด่างของดิน (pH) อยู่ในช่วง 5.0-7.0 มีหน้าดินลึกระบายน้ำ

2) แหล่งน้ำ น้ำเป็นสิ่งจำเป็นต่อกาเจริญเติบโตของลำไย การผลิตลำไยเพื่อให้ได้คุณภาพต้องมีน้ำในปริมาณที่เพียงพอตลอดฤดูกาล

3) สภาพภูมิอากาศ ปัจจัยสภาพภูมิอากาศที่มีบทบาทสำคัญต่อการเจริญเติบโตของลำไย ได้แก่

3.1) อุณหภูมิ โดยทั่วไปลำไยต้องการอากาศค่อนข้างเย็น อุณหภูมิที่สามารถเจริญเติบโตได้อยู่ระหว่าง 4-30 องศาเซลเซียส และต้องการอุณหภูมิต่ำ 10-22 องศาเซลเซียส ในช่วงฤดูหนาวเดือนพฤศจิกายนถึงมกราคม เพื่อสร้างตาดอก ซึ่งในปีที่มีอากาศเย็นระยะเวลานานโดยไม่มีอากาศอุ่นแทรก ลำไยจะออกดอกติดผลดี แต่ถ้ามีอุณหภูมิไม่ต่ำพอ ต้นลำไยจะออกดอกน้อยหรือไม่ออกดอก

3.2) แสง การเจริญเติบโตของลำไยจาเป็นต้องได้รับแสงอย่างเพียงพอ ดังนั้นการปลุกลำไยจึงควรปลูกในที่โล่ง ในสภาพพื้นที่ที่มีปริมาณแสงน้อยซึ่งอาจเกิดจากการบังแสงของเมฆ หรือเกิดฝนตกติดต่อกันหลายวัน มักทำให้ต้นลำไยชะงักการเจริญเติบโต ส่วนในสภาพที่มีความเข้มแสงสูงมักเกิดปัญหาทำให้ผิวของผลลำไยเป็นสีน้ำตาลเข้มจำหน่ายได้ราคาตกต่ำ

3.3) ปริมาณน้ำฝนและความชื้นสัมพัทธ์ แหล่งปลูกลำไยควรมีปริมาณน้ำฝนอยู่ในช่วงประมาณ1000 – 200 มิลลิเมตรต่อปี และควรมีการกระจายของฝนประมาณ100-150 วันต่อปีในแหล่งปลูกที่มีปริมาณฝนตกน้อย ควรจัดหาแหล่งน้ำและระบบชลประทานให้เพียงพอและเหมาะสม

3.4) ระดับความสูงของพื้นที่ ลำไยสามารถปลูกได้ดีในที่ราบลุ่มจนถึงพื้นที่สูงกว่าระดับน้ำทะเล 1000 เมตร

4) รูปแบบการปลูกลำไย รูปแบบการปลุกลำไยที่นิยมมี 3 แบบ (พาวิน และคณะ, 2547) คือ

4.1) การปลูกระยะห่าง เป็นวิธีที่นิยมมากตั้งแต่ในอดีตและปัจจุบัน การปลูกลำไยต้องการให้ต้นลำไยมีเจริญเติบโตขยายขนาดของทรงพุ่มเต็มที่ รูปแบบการปลุกมีทั้งสี่เหลี่ยมจัตุรัส และแบบสี่เหลี่ยมผืนผ้า โดยกำหนดให้ระยะห่างระหว่างแถวและระยะห่างระหว่างต้นเกิน 8 เมตร เช่น 8x8 10x10 12x12 8x10 และ 10x12 เมตร ต้นลำไยมักมีทรงพุ่มขนาดสูงใหญ่ ปริมาณผลผลิตสูง แต่จำนวนต้นต่อไร่น้อยมักประสบปัญหาการจัดการและต้นลำไยโค่นล้มง่ายโดยเฉพาะเมือเกิดพายุลมแรง

4.2) การปลูกระยะชิด เป็นการใช้พื้นที่ให้เกิดประโยชน์สูงสุด แต่การปลูกระยะชิดต้องมีการตัดแต่งกิ่งเอควบคุมทรงพุ่มและการใช้สารโพแทสเซียมคลอเรต กระตุ้นให้มีการออกดอก การปลูกลำไยระยะชิดเป็นรูปแบบการปลูกที่ได้จานวนต้นต่อไร่สูง ในประเทศไทยมีการสร้างสวนลำไยระยะชิดยังไม่แพร่หลาย อาจเนื่องมาจากมีบทเรียนจากการปลูกลำไยระยะชิด ที่ไม่ประสบความสำเร็จในอดีต

การควบคุมทรงพุ่มทำได้ยากเพราะต้นลำไยที่ตัดแต่งกิ่งมักออดอกปีเว้นปี ซึ่งรูปแบบการปลุกลำไยระยะชิดมีหลายๆ แบบ ดังนี้

4.2.1) การปลูกระยะชิดแบบแถวเดี่ยว เป็นรูปแบบการปลูกคล้ายระบบการปลูกห่างแต่มีระยะปลูกที่แคบกว่า เช่น แบบสี่เหลี่ยมจัตุรัส ระยะปลูก 4x4 5x5 เมตร หรือ แบบสี่เหลี่ยมผืนผ้า ระยะปลูก 3x6 4x6 เมตร ซึ่งสามารถนาเครื่องจักรเข้าไปปฏิบัติงานในสวนได้สะดวกกว่าแบบสี่เหลี่ยมจัตุรัส

4.2.2) การปลูกระยะชิดแบบแถวคู่ เป็นระบบการปลูกที่กำหนดให้แถวอยู่ชิดกันหนึ่งคู่สลับกับแถวห่างเพื่อการปฏิบัติงานงานโดยเครื่องจักร เป็นระบบที่เพิ่มจานวนต้นต่อไร่มากขึ้นและมีพื้นที่การให้ผลผลิตเพิ่มมากขึ้น แต่ในลักษณะสภาพแวดล้อมที่อากาศร้อนชื้นอาจทำให้มีการระบาดของโรคและแมลงศัตรูลำไยมาก

4.2.3) การปลูกระยะชิดแบบกลุ่ม เป็นระบบการปลูกลำไยรวมกันให้เกิดเป็นกลุ่มโดยอาศัยเทคนิคการตัดแต่งกิ่งควบคุมทรงพุ่ม เป็นการเพิ่มพื้นที่ของการให้ผลผลิตลำไย

4.3) ระบบคอนทัวร์หรือระบบแนวระดับ เป็นระบบการปลูกลำไยที่ช่วยป้องกันและลดอัตราการชะล้าง หรือการพังทลายของดินในพื้นที่ที่มีความลาดชัน ปกติระบบการปลุกนี้จะให้เมื่อพื้นที่ปลูกมีความลาดชันเกิน 3 เปอร์เซ็นต์ หมายถึงในทุกระยะทาง 100 เมตร จะมีระดับความสูงขึ้นหรือต่ำลง 3 เมตรขึ้นไปต้องทาการปลูกตามแนวระดับ การเตรียมพื้นที่ปลูกต้องมีการทำระดับหรือขั้นบันได ตามระดับความสูงของพื้นที่ซึ่งการปลูกแบบนี้มีความยุ่งยากต่อการปฏิบัติงานในสวนมากกว่าวิธีอื่น

- **การปฏิบัติดูแลรักษาลำไย**

1) แทงช่อดอก

1.1) เริ่มให้น้ำอย่างสม่ำเสมอ

1.2) เฝ้าระวังโรคแมลงศัตรูในช่วงแทงช่อดอก เช่นหนอนม้วนใบซึ่งจะม้วนใบและช่อดอก แมลงค่อมทองและมวนลำไย ถ้าพบเห็นให้รีบป้องกันกำจัด

2) ดอกบาน

2.1) งดการพ่นสารเคมีเพราะจะเป็นอันตรายต่อแมลงที่จะช่วยผสมเกสร

2.2) สำรวจโรคและแมลง ช่วงนี้อาจมีการเข้าทำลายของ หนอน และเพลี้ยไพเข้าทำลาย ถ้าพบเห็นให้รีบป้องกันกำจัด

2.3) ควรให้น้ำอย่างสม่ำเสมอ

3) ปรับปรุงคุณภาพผลผลิต

ตัดแต่งช่อผลในกรณีที่ต้นลำไยติดผลดกต้องตัดช่อผลให้เหลือประมาณ 40-60ผลต่อช่อ และตัดช่อผลเล็กออกบ้างจะทำให้ผลมีขนาดใหญ่ขึ้น

4) ตัดแต่งกิ่ง

ตัดแต่งกิ่งให้เป็นทรงพุ่มโปร่งโดยเลือกรูปทรงให้ เหมาะสมกับสภาพต้นและสวน นำใบลำไยที่ตัดแต่งกิ่งคลุมไว้ที่โคนต้น

5) การดูแลรักษาช่วงติดผล-เก็บเกี่ยว

5.1) เฝ้าระวังการทำลายของโรค และแมลง

5.2) ควรให้น้ำอย่างสม่ำเสมอ

5.3) ให้ปุ๋ยสูตร 46-0-0 ร่วมกับ 15-15-15 และ 0-0-60 อัตราขึ้นอยู่กับขนาดทรงพุ่ม ปริมาณการออกดอกติดผล หรือผลผลิตและผลการวิเคราะห์ดิน

6) เก็บเกี่ยวผลผลิต

6.1) สังเกตจากผลลำไยมีผิวเรียบ ผลมีขนาดใหญ่ ไม่ควรหักหรือตัดกิ่งลึกเกิน 30 เซนติเมตร

6.2) ลำเลียงผลผลิตเข้าในที่ร่มเพื่อคัดเกรดและบรรจุลงตะกร้าเพื่อเตรียมจำหน่าย

7) การบำรุงใบและสะสมอาหาร

7.1) ให้ปุ๋ยอินทรีย์ 10-20 กิโลกรัมต่อต้น

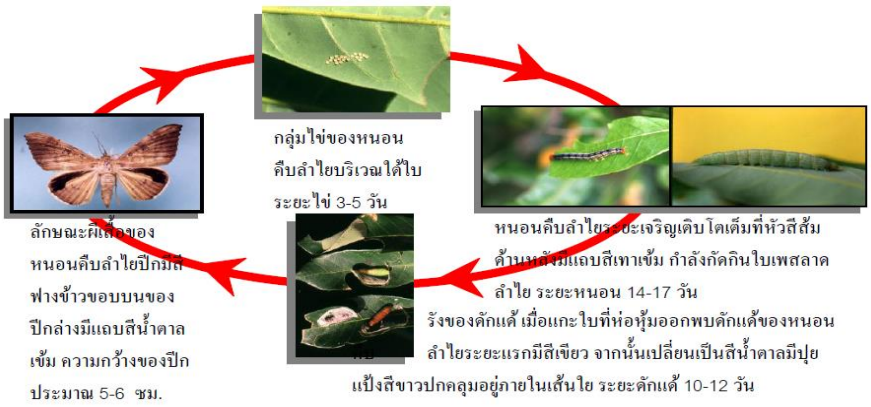
7.2) ให้ปุ๋ยเคมีสูตร 46-0-0 และ 15-15-15 และ 0-0-60 (แตกใบครั้งที่หนึ่ง)

7.3) ให้ปุ๋ยเคมีสูตร 46-0-0 และ 15-15-15 และ 0-0-60 (แตกใบครั้งที่สอง)

7.4) เพิ่มเปอร์เซ็นต์การออกดอกโดยใช้สารโพแทสเซียมคลอเรต

6) โรคและแมลงศัตรูสำคัญของลำไย

แมลงที่พบระบาดเป็นประจำในช่วงที่ลำไยแตกใบอ่อน คือหนอนคืบ ลำไย (*Oxyodes scrobiculatus)* และหนอนคืบเขียวกินใบ (*Thalasodes* spp.) แมลงค่อมทอง และอาการพุ่มไม้กวาด จากภาพที่ ก.2 แสดงวงจรชีวิตของนอนคืบ ภาพที่ ก.3 แสดงตัวอย่างแมลงค่อมทอง และภาพที่ ก.3 แสดงใบม้วนหงิก (จริยา วิสิทธิ์พานิช และชาตรี สิทธิกุล, 2548)



ภาพที่ ก.2 วงจรชีวิตหนอนคืบ

ภาพที่ ก.3 แมลงค่อมทอง

ภาพที่ ก.4 อาการใบม้วนหงิกบนช่อใบอ่อนและดอกลำไยพันธุ์ดอ

- **การป้องกันและรักษา**

1) ระยะแตกใบอ่อน หนอนคืบกินใบ แมลงค่อมทอง (ฉีดพ่นด้วยคาร์บาริล 85% WP อัตรา 45 กรัมต่อน้ำ 20 ลิตร)

2) ระยะออกดอก หนอนม้วนใบ แมลงค่อมทอง มวนลำไย (ฉีดพ่นด้วยแลมป์ดาไซฮาโลทริน 25 % EC อัตรา 10 ซีซี ต่อน้ำ 20 ลิตร หรือ คาร์บาริล อัตรา 45 กรัม ต่อ น้ำ 20 ลิตร) เพลี้ยไฟ (ฉีดพ่นด้วยอิมิดาโคลพริด 8 ซีซี. ต่อน้ำ 20 ลิตร) โรคพุ่มไม้กวาด ตัดกิ่งที่เป็นโรคทิ้ง

3) ระยะติดผล เพลี้ยหอย เพลี้ยแป้ง (พ่นด้วยคลอไพริฟอส 30 ซีซี. ผสมปิโตรเลียมออย 30 ซีซี. น้ำ 20 ลิตร)

4) ระยะก่อนเก็บเกี่ยว โรคผลแตกผลลาย เพลี้ยหอย เพลี้ยแป้ง (ฉีดพ่นด้วยคาร์เบนดาซิม หรือ เบนเลท หรือ ทีบูโคนาโซล)

- **การจัดการหลังการเก็บเกี่ยว**

หลังจากบรรจุลำไยใส่ตะกร้าแล้ว ลำไยที่จะส่งออกมักจะนำไปรมด้วยแก๊สซัลเฟอร์ไดออกไซด์ (SO_2) เพื่อควบคุมป้องกันโรคหลังการเก็บเกี่ยวและฟอกสีผิวทาให้มีผิวสีทองสวยตรงตามความต้องการของตลาด การรมลำไยสดด้วยซัลเฟอร์ไดออกไซด์ ควรต้องทำในห้องรมที่สร้างได้ตามมาตรฐานของสถาบันอาหารและคำนวณปริมาณผงกามะถันตามที่แนะนำ และรมไว้ในระยะเวลาที่กำหนดอย่างเคร่งครัดเพื่อป้องกันการตกค้างที่มากจนเกินไปมาตรฐานของ SO_2 บนเปลือกและเนื้อของลำไยสด โดยตัวอย่างของปริมาณลำไย ผงกามะถันและเวลาในการรมแสดงไว้ในตารางที่ 2.1 (ธีรนุช เจริญกิจ และพาวิน มะโนชัย, 2548)

ตารางที่ ก.1 ปริมาณของลำไย ปริมาณของผงกำมะถันที่ใช้และระยะเวลาที่ใช้รมลำไยสด (ธีรนุช เจริญกิจ และพาวิน มะโนชัย, 2548)

ปริมาณลำไยสด (กิโลกรัม)	ปริมาณกำมะถันผง (กรัม)	ระยะเวลาที่เผา (นาที)
350	3.0	10
1,500	13.5	20
2,500	24.0	35
3,800	36.0	45
5,000	47.0	35

- **สรรพคุณทางยา**

ใบ เป็นใบสด มีรสจืดและชุ่ม สุขุม เป็นยาแก้โรคมาลาเรีย ริดสีดวงทวาร ฝีหัวขาด และแก้ไข้หวัด โดยนำเอาต้มน้ำกิน

ดอก ใช้ดอกสด หรือตากแห้งเก็บไว้ใช้ เป็นยาแก้โรคเกี่ยวกับหนองทั้งหลาย โดยใช้ใบสดประมาณ 5-30 กรัมต้มน้ำกิน

เมล็ด ต้มหรือบดเป็นผงกินจะมีรสฝาด ใช้ภายนอกจะรักษากลากเกลื้อน แผลมีหนอง แก้ปวดสมานแผล ใช้ห้ามเลือด

รากหรือเปลือกราก ต้มน้ำกินหรือเคี่ยวให้ข้นผสมกิน มีรสฝาด แก้สตรีตกขาวมากผิดปกติ ขับพยาธิเส้นด้าย

เปลือกผล ใช้ที่แห้งนำมาต้มน้ำกิน แก้อาการวิงเวียนศรีษะ และอ่อนเพลีย ทำให้สดชื่น จะมีรสขม หรือใช้ทาภายนอก โดยเผาเป็นเถ้าหรือบดเป็นผงโรยแผลไฟไหม้ น้ำร้อนรวก

เนื้อหุ้มเมล็ด นำมาต้มน้ำกินหรือแช่เหล้า เป็นยาบำรุงม้าม เลือด และหัวใจ บำรุงร่างกาย สงบประสาท แก้อ่อนเพลียจากการทำงานหนัก ลืมง่าย นอนไม่หลับ ประสาทอ่อน หรือจะบดเป็นผงผสมกับยาเม็ดกิก็ได้

ข้อห้ามใช้ คนที่มีอาการเจ็บคอ หรือไอ มีเสมหะ หรือเป็นแผลอักเสบจนมีหนอง ไม่ควรกินเนื้อของผลลำไย

ผลการวิจัยของ ยุทธนา สุดเจริญ (2548) พบว่าเมล็ดลำไย มีสารต้านอนุมูลอิสระ หนึ่งในสารสำคัญที่ได้จากการสกัดแยกของเมล็ดลำไย คือ อีลาจิก แอซิค (Ellagic Acid) ซึ่งเป็นสาร ป้องกันการก่อมะเร็ง(Anticarcinogenic Agent) และเป็นสารที่ป้องกันการก่อกลายพันธุ์ (Antimutagenic Compound) อีกทั้งยังพบในปริมาณสูงมากเมื่อเทียบกับพืชจำพวกราสเบอรี่ ในผลการวิจัยจากต่างประเทศ

(2) ทุเรียน

- **ลักษณะทางพฤกษศาสตร์ (เดโชดม ภัทรศัย, 2543)**

ชื่อภาษาไทย ทุเรียน

ชื่อภาษาอังกฤษ Durion

ชื่อวิทยาศาสตร์ *Durio zibethinus.*

วงศ์ Malvaceae (Bombacaceae)

ทุเรียนเป็นต้นไม้ยืนต้น มีลำต้นใหญ่ ทรงพุ่มใหญ่ ลำต้นจะมีเปลือกสีเทาแก่ที่รอยหยาบ เป็นทางยาว กิ่งก้านจะแตกออกจากต้น โดยสลับทิศทางกันรอบต้น มีลักษณะคดโค้งบ้างหยาบบ้างขึ้นอยู่กับสายพันธุ์ ใบ เป็นใบเดี่ยวและแตกออกจากตา ของกิ่งในลักษณะทแยงตรงกันข้ามกับกิ่งบนต้น ดอกทั้งดอกตูมที่มีลักษณะรีๆ คล้ายรูปไข่มีลักษณะกลมและดอกสมบูรณ์เพศ ผล ผลอ่อนมีสีเขียวสด ผลแก่

จะเปลี่ยนแปลงเป็นสีน้ำตาลหรือเขียวแกมเทา ปลายของหนามทุเรียนจะแห้ง แสดงให้เห็นว่าพูเกิดอาการหลวมตัวทำให้มีช่องว่างระหว่าเนื้อกับเปลือก ก้านผลจะมีลักษณะบวมโต โคนหนามผลแก่จะมองเห็นได้เท่าหัวเข็มหมุด ไม่เหมือนผลอ่อนที่ไม่มีหนาม

- **การเลือกพื้นที่ปลูก (วันทนา บัวทรัพย์ และ มนตรี วงศ์รักษ์พานิช , 2548)**

1) แหล่งน้ำต้องมีแหล่งน้ำจืดให้ต้นทุเรียนได้เพียงพอตลอดปี

2) อุณหภูมิและความชื้น ทุเรียนชอบอากาศร้อนชื้นอุณหภูมิที่เหมาะสมอยู่ในช่วงประมาณ 25-30 องศาเซลเซียส ความชื้นสัมพัทธ์ในอากาศประมาณ 75-85% ถ้าปลูกในพื้นที่ที่มีอากาศแห้งแล้ง มีอากาศร้อนจัดเย็น และมีลมแรง จะพบปัญหาใบไหม้หรือใบร่วง ต้นทุเรียนไม่เจริญเติบโตหรือเติบโตช้าให้ผลผลิตช้า และน้อยไม่คุ้มต่อการลงทุน

3) สภาพดิน ควรเป็นดินร่วน ดินร่วนปนทราย ดินเหนียวปนทราย ที่มีการระบายน้ำดีและมีหน้าดินลึก เพราะทุเรียนเป็นพืชที่อ่อนแอต่อสภาพน้ำขัง ความเป็นกรดด่างของดินอยู่ระหว่าง 5.5-6.5 ถ้าจำ เป็นต้องปลูกทุเรียนในสภาพดินทราย จำ เป็นต้องนำ หน้าดินจากแหล่งอื่นมาเสริม ต้องใส่ปุ๋ยคอกและต้องดูแลเรื่องการให้น้ำมากเป็นพิเศษ แหล่งน้ำต้องเพียงพอ

- **พันธุ์ทุเรียน**

ที่มา : วันทนา บัวทรัพย์ และ มนตรี วงศ์รักษ์พานิช, 2548

ภาพที่ ก.5 พันธุ์ทุเรียน

1) ชะนี

ข้อดี ทนทานต่อโรครากเน่าโคนเน่าพอสมควร ออกดอกง่าย เนื้อแห้ง รสดี สีสวย

ข้อเสีย ออกดอกดกแต่ติดผลยาก เป็นไส้ซึมง่าย และอ่อนแอต่อโรคใบติด

2) หมอนทอง

ข้อดี ราคาสูงกว่าพันธุ์อื่น ติดผลดีมากน้ำหนักผลดี เนื้อมากเมล็ดลีบ มีกลิ่นน้อย เนื้อละเอียดแห้ง ไม่เละ ผลสุกแล้วเก็บไว้ได้นาน และไม่ค่อยเป็นไส้ซึม

ข้อเสีย อ่อนแอต่อโรครากเน่าโคนเน่า

3.3) ก้านยาว

ข้อดี ติดผลดี ราคาค่อนข้างดี น้ำหนักผลดี

ข้อเสีย ค่อนข้างอ่อนแอต่อโรค รากเน่า โคนเน่า เปลือกหนา เนื้อน้อย เป็นไส้ซึม ค่อนข้างง่าย ผลสุกเก็บไว้ได้ไม่นาน ก้นผล แตกง่ายและอายุการให้ผลช้า

3.4) กระตูม

ข้อดี ไม่มีปัญหาไส้ซึมเพราะเป็นพันธุ์เบาเก็บเกี่ยวก่อนฝนตกชุก ออกดอกเร็ว ผลแก่เร็วจึงขายได้ราคาดีในช่วงต้นฤดู ผลดก ติดผลง่าย อายุการให้ผลเร็ว

ข้อเสีย อ่อนแอต่อโรครากเน่าโคนเน่า

- **การปลูก**

ฤดูปลูก ถ้ามีการจัดระบบการให้น้ำอย่างมีประสิทธิภาพ สามารถดูแลให้น้ำกับต้นทุเรียนได้สม่ำเสมอช่วงหลังปลูก ควรปลูกตั้งแต่เดือนมีนาคม-เมษายน แต่ถ้าหากจัดระบบน้ำไม่ทันหรือยังไม่อาจดูแลเรื่องน้ำได้ ควรปลูกต้นฤดูฝนเตรียมพื้นที่การปลูกทุเรียน

1) ไถ ขุดตอ ขุดรากไม้เก่า ออกจากแปลง

1.1) พื้นที่ดอนไม่มีปัญหาน้ำท่วมขัง : ไถกำจัดวัชพืชอย่างเดียว

1.2) พื้นที่ดอน มีแอ่งที่ลุ่มน้ำขัง : ไถปรับสภาพพื้นที่ให้เรียบ

1.3) พื้นที่ลุ่มหรือต่ำมีน้ำท่วมขัง : ทำทางระบายน้ำหรือยกร่อง

2) กำหนดระยะปลูก

ระยะระหว่างต้นและระยะระหว่างแถวด้านละ 9 เมตร ปลูกได้ไร่ละ 20 ต้น การทำสวนขนาดใหญ่ ควรขยายระยะระหว่างแถวให้กว้างขึ้น เพื่อสะดวกต่อการนำ เครื่องจักรกลต่างๆ ไปทำงานในระหว่างแถว

3) วางแนวและปักไม้ตามระยะปลูกที่กำหนด

วางแนวกำหนดแถวปลูกโดยคำนึงว่า แนวปลูกขวางความลาดเทของพื้นที่ หรืออาจกำหนดในแนวตั้งฉากกับถนน หรือกำหนดแถวปลูกไปในแนวทิศตะวันออก ตะวันตก และถ้ามีการจัด

วางระบบน้ำ ต้องพิจารณาแนวทางจัดวางท่อในสวนด้วย จากนั้นจึงปักไม้ตามระยะที่กำหนดเพื่อขุดหลุมปลูกต่อไป วิธีการปลูกทุเรียนทำ ได้ 2 ลักษณะ

3.1) วิธีการขุดหลุมปลูก เหมาะกับสวนที่ไม่มีการวางระบบน้ำ

- ขุดหลุมมีขนาดกว้างยาว และลึกด้านละ 50 เซนติเมตร

- ผสมปุ๋ยคอกเก่าประมาณ 5 กิโลกรัม และปุ๋ยหินฟอสเฟตครึ่งกิโลกรัม คลุกเคล้ากับดินที่ขุดขึ้นมา กลบกลับคืนไปในหลุมสูงประมาณ 2 ใน 3 ของหลุม

- เตรียมต้นกล้าที่แข็งแรงสมบูรณ์ ไม่เป็นโรค ไม่มีแมลงทำ ลาย และมีใบยอดคู่สุดท้ายแก่ระบบรากแผ่กระจายดี ไม่ขดม้วนงออยู่ก้นถุง

- ใช้มีดกรีดก้นถุงออก ถ้าพบรากขดงออยู่ก้นถุงให้ตัดออก

- วางถุงต้นกล้าที่ตัดก้นถุงออกแล้ววางลงตรงกลางหลุม จัดให้ตรงแนวกับต้นอื่นๆ พร้อมทั้งปรับระดับสูงต่ำของต้นทุเรียนให้รอยต่อระหว่างรากกับลำ ต้นหรือระดับดินปากถุงเดิมสูงกว่าระดับดินปากหลุมเล็กน้อย

- ใช้มีดกรีดด้านข้างถุงจากล่างขึ้นบนทั้งสองด้าน

- ดึงถุงพลาสติกออก ระมัดระวังอย่าให้ดินในถุงแตก

- กลบดินที่เหลือลงไปในหลุมอย่ากลบดินสูงถึงรอยเสียบยอด หรือรอยทาบ

- ปักไม้หลักข้างต้นทุเรียนที่ปลูกแล้ว พร้อมทั้งผูกเชือกยึดไว้เพื่อป้องกันลมพัดโยก

- กดดินบริเวณโคนต้น หาวัสดุคลุมโคนต้นแล้วรดน้ำตามให้โชก

- จัดทำ ร่มเงาให้ต้นทุเรียนที่เพิ่งปลูก โดยใช้ทางมะพร้าว ทางจาก แผงหญ้าคา ทางระกำหรือตาข่ายพรางแสง เมื่อทุเรียนตั้งตัวดีแล้วควรปลดออก หรืออาจปลูกไม้เพื่อให้ร่มเงา เช่น กล้วยก็จะช่วยเป็นร่มเงาและเพิ่มความชื้นในสวนทุเรียนได้เป็นอย่างดี โดยเฉพาะในช่วงฤดูแล้งที่อากาศแห้งและมีแสงแดดจัด

- แกะผ้าพลาสติกที่พันรอยเสียบยอดหรือทาบออกเมื่อปลูกไปแล้วประมาณ 1-2 เดือน

3.2) วิธีการปลูกแบบไม่ขุดหลุม เหมาะกับสวนที่จัดวางระบบน้ำ มีข้อดีคือ ประหยัดแรงงานค่าใช้จ่ายในการขุดหลุม ดินระบายน้ำและอากาศดี รากเจริญเร็ว

- โรยปุ๋ยหินฟอสเฟต 500 กรัม หรือประมาณหนึ่งกระป๋องนมครึ่ง ตรงตำแหน่งที่ต้องการปลูกกลบดินบางๆ

- นำต้นพันธุ์มาวาง แล้วถากดินข้าง ๆ ขึ้นมาพูนกลบ แต่ถ้าหากเป็นดินร่วนปนทราย ดิน ทราย ดินจะไม่เกาะตัวกัน ควรใช้วิธีขุดหลุมปลูก หรือจะใช้วิธีดัดแปลง

- วิธีดัดแปลง คือ นำ หน้าดินจากแหล่งอื่นมากองตรงตำแหน่งที่จะปลูก กองดินควรมีขนาดเส้นผ่าศูนย์กลางประมาณ 1 เมตร สูง 15 เซนติเมตร แหวกกลางกองดินโรยปุ๋ยหินฟอสเฟตในช่องที่แหวกไว้ กลบดินบางๆ วางต้นพันธุ์ดีลงตรงช่องที่แหวกไว้กลบดินทับ

- การแกะถุงออก ต้องระมัดระวังอย่าให้ดินแตกอาจทำได้โดยกรีดถุงออกก่อนแล้วนำไปวางในตำแหน่งที่ปลูก กรีดถุงพลาสติกให้ขาดจากล่างขึ้นบน แล้วจึงค่อยๆ ดึงถุงพลาสติกออกเบาๆ

- ระมัดระวังอย่ากลบดินให้สูงถึงรอยเสียบยอดหรือรอยทาบ

- หาวัสดุคลุมโคน และจัดทำ ร่มเงาให้กับต้นทุเรียนเหมือนการปลูกโดยวิธีขุดหลุม

- **การปฏิบัติดูแลรักษาทุเรียน**

1) การปฏิบัติดูแลทุเรียนในช่วงก่อนให้ผลผลิต เป็นสิ่งสำคัญที่จะทำ ให้ทุเรียนเจริญเติบโตอย่างรวดเร็ว และให้ผลผลิตได้เร็วขึ้น

1.1) ในระหว่างรอทุเรียนให้ผลผลิต ในช่วงแรกควรปลูกพืชแซมเสริมรายได้ โดยเลือกพืชให้ตรงกับความต้องการของตลาด

1.2) เมื่อตรวจพบทุเรียนตายหลังปลูกให้ทำ การปลูกซ่อม

1.3) การให้น้ำ ช่วงเวลาหลังจากปลูกจะตรงกับฤดูฝน ถ้ามีฝนตกหนักควรทำ ทางระบายน้ำ และตรวจดูบริเวณหลุมปลูก ถ้าดินยุบตัวเป็นแอ่งมีน้ำขังต้องพูนดินเพิ่ม ถ้าฝนทิ้งช่วง ควรรดน้ำให้ดินมีความชื้นอยู่เสมอ ในปีต่อๆ ไป ควรดูแลรดน้ำให้ต้นไม้ผลอย่างสม่ำเสมอ และในช่วงฤดูแล้งควรใช้วัสดุคลุมดิน เพื่อช่วยรักษาความชื้นในดิน เช่น ฟางข้าว หญ้าแห้ง

1.4) การตัดแต่งกิ่ง ปีที่ 1-2 ไม่ควรตัดแต่ง ปล่อยให้ต้นทุเรียนเจริญเติบโตอย่างเต็มที่ ปีต่อ ๆ ไป ตัดแต่งกิ่งแห้ง กิ่งแขนง กิ่งกระโดงในทรงพุ่ม กิ่งเป็นโรคออก เลี้ยงกิ่งแขนงที่สมบูรณ์ที่อยู่ในแนวขนานกับพื้น (กิ่งมุมกว้าง) ไว้ในปริมาณและทิศทางเหมาะสม โดยให้กิ่งล่างสุดอยู่สูงจากพื้นดินประมาณ 80-100 เซนติเมตร

1.5) การป้องกันกำจัด ช่วงแตกใบอ่อน : ควรป้องกันกำ จัดโรคใบติด เพลี้ยไก่แจ้ เพลี้ยไฟ ไรแดง ช่วงฤดูฝน:ป้องกันกำ จัดโรครากเน่าโคนเน่าและควบคุมวัชพืชโดยการปลูกพืชคลุมดินและอาจจะกำ จัดโดยใช้แรงงานขุด ถาก ถอน ตัด พยายามหลีกเลี่ยงการใช้สารเคมีเพราะต้นทุเรียนยังเล็กอยู่ละอองสารเคมีอาจจะไปทำลายต้นทุเรียน

1.6) การทำร่มเงาในช่วงฤดูแล้ง แสงแดดจัดมาก ทำ ให้ทุเรียนใบไหม้ได้ ควรทำ ร่มเงาให้

1.7) การใส่ปุ๋ยควรใส่ปุ๋ยหลังจากตัดแต่งกิ่ง โดยใส่ปุ๋ยพร้อมกับการทำ โคน คือ ถากวัชพืชบริเวณใต้ทรงพุ่ม หว่านปุ๋ยและพรวนดินนอกชายพุ่มขึ้นมากลบใต้ทรงพุ่มให้มีลักษณะเป็นหลังเต่า และขยายขนาดของเนินดินให้กว้างขึ้นตามขนาดของทรงพุ่มหรือจะใส่ปุ๋ย โดยวิธีใช้ไม้ปลายแหลมแทงดินเป็นรูหยอดปุ๋ยใส่และปิดหลุมเป็นระยะให้ทั่วบริเวณใต้ทรงพุ่มวิธีหลังนี้แม้จะเปลืองแรงงานแต่ช่วยลดการสูญเสียของปุ๋ยจากการระเหย หรือถูกน้ำชะพา

1.8) หว่านปุ๋ยคอกก่อนและตามด้วยปุ๋ยเคมี

1.9) ควรใส่ปุ๋ยในบริเวณใต้ทรงพุ่มโดยรอบ และให้ห่างจากโคนต้นประมาณ 20-30 เซนติเมตรขึ้นไป ขึ้นกับขนาดทรงพุ่มปริมาณและเวลาใส่ปุ๋ย ปีที่ 1 : ใส่ปุ๋ยและทำโคน 4 ครั้ง (เดือนเว้นเดือน) ครั้งที่ 1-3 ใส่ปุ๋ยคอก 5 กิโลกรัมต่อต้น (ประมาณ 1 ปีบ) ครั้งที่ 4 - ใส่ปุ๋ยคอก 5 กิโลกรัมต่อต้น (ประมาณ 1 ปีบ)

1.10) ใส่ปุ๋ยเคมีสูตร 15-15-15 หรือ16-16-16 ประมาณ150-200 กรัมต่อต้น (ครึ่งกระป๋องนมข้น) ปีต่อ ๆ ไป(ระยะที่ทุเรียนยังไม่ให้ผลผลิต):ใส่ปุ๋ยและทำ โคน 2 ครั้ง (ต้นฤดูฝนและปลายฤดูฝน)

ครั้งที่ 1 (ต้นฝน) ใส่ปุ๋ยเคมีสูตร 15-15-15 หรือ 16-16-16ประมาณครึ่งถึง 3 กิโลกรัมต่อต้น

ครั้งที่ 2 (ปลายฝน) ใส่ปุ๋ยคอก 15-50 กิโลกรัมต่อต้น (ประมาณ 3-10 ปีบ) ใส่ปุ๋ยเคมีสูตร 15-15-15 หรือ 16-16-16 ประมาณครึ่งถึง 3 กิโลกรัมต่อต้น ปริมาณปุ๋ยเคมีที่ใส่ในแต่ละครั้งขึ้นกับขนาดของทรงพุ่ม โดยยึดหลักว่า วัดจากโคนต้นมายังชายพุ่มเป็นเมตรได้เท่าไร คือ จำนวนปุ๋ยเคมีที่ใส่เป็นกิโลกรัม เช่น ระยะจากโคนต้นถึงชายพุ่ม 1 เมตร ใส่ปุ๋ย 1 กิโลกรัม ระยะจากโคนต้นถึงชายพุ่ม 2 เมตร ใส่ปุ๋ย 2 กิโลกรัม ระยะจากโคนต้นถึงชายพุ่ม 2 เมตรครึ่ง ใส่ปุ๋ย 2 กิโลกรัมครึ่ง

2) การปฏิบัติดูแลทุเรียนในช่วงให้ผลแล้ว เป็นสิ่งสำคัญที่จะทำ ให้ทุเรียนออกดอกติดผลมากและให้ผลผลิต คุณภาพดี การเตรียมต้นให้พร้อมที่จะออกดอก คือการเตรียมให้ต้นทุเรียนมีความสมบูรณ์มีอาหารสะสมเพียงพอ เมื่อทุเรียนใบแก่ทั้งต้น และสภาพแวดล้อมเหมาะสม ฝนแล้ง ดิน มีความชื้นต่ำ อากาศเย็นลงเล็กน้อยทุเรียนก็จะออกดอก ขั้นตอนต่างๆ จะต้องรีบดำเนินการภายหลังจากเก็บเกี่ยวผลผลิต ดังนี้

2.1) การตัดแต่งกิ่ง หลังเก็บเกี่ยวให้รีบตัดแต่งกิ่งแห้ง กิ่งเป็นโรค กิ่งแขนง ด้านในทรงพุ่มออกโดยเร็ว ทารอยแผลที่ตัดด้วยสารเคมีป้องกันกำ จัด เชื้อรา หรือปูนแดงกินกับหมาก

2.2) หลังตัดแต่งกิ่ง ให้กำ จัดวัชพืชและใส่ปุ๋ยทันที โดยใส่ปุ๋ยคอก 15-50 กิโลกรัมต่อต้น (ประมาณ 3-10 ปีบ) และปุ๋ยเคมีสูตร 15-15-15 หรือ 16-16-16 ในอัตรา 3-5 กก. ต่อต้น

2.3) ในช่วงฤดูฝน ถ้าฝนตกหนัก จัดการระบายน้ำออกจากแปลงปลูก ถ้าฝนทิ้งช่วงให้รดน้ำแก่ต้นทุเรียน ควบคุมวัชพืช โดยการตัดและ หรือใช้สารเคมี ป้องกันกำจัดโรคแมลง เช่น โรครากเน่าโคนเน่า โรคใบติด โรคแอนแทรกโนส เพลี้ยไก่แจ้ ไรแดงและเพลี้ยไฟ

2.4) ในช่วงปลายฤดูฝน เมื่อฝนทิ้งช่วง ให้ใส่ปุ๋ยเคมีสูตร 8-24-24, 9-24-24 หรือ 12-24-12 2-3 กก.ต่อต้นเพื่อช่วยในการออกดอก ให้กำ จัดวัชพืชใต้ทรงพุ่ม กวาดเศษหญ้า และใบทุเรียนออกจากโคนต้น เพื่อให้ดินแห้งเร็วขึ้น งดการให้น้ำ10-14 วัน เมื่อสังเกตเห็นใบทุเรียนเริ่มลดลงต้องเริ่มให้น้ำทีละน้อยเพื่อกระตุ้นให้ตามดอกเจริญอย่าปล่อยให้ขาดน้ำนานจนใบเหลืองใบตก เพราะตาดอกจะไม่เจริญ และระวังอย่าให้น้ำมากเกินไป เพราะช่อดอกอาจเปลี่ยนเป็นใบได้ ซึ่งวิธีให้น้ำที่เหมาะสม คือ ให้น้ำแบบโชย ๆ แล้วเว้นระยะ สังเกตอาการของใบและดอก เมื่อเห็นดอกระยะไข่ปลามากพอแล้ว ก็เพิ่มปริมาณให้มากขึ้นเรื่อยๆ จนสู่สภาวะปกติ

- **โรคทุเรียน (ศุภลักษณ์ กลับน่วม และอัญชลี พัดมีเทศ, 2547)**

1) โรครากเน่าและโคนเน่า

เกิดจาก เชื้อราเจริญเติบโตเข้าไปทำลายทุเรียนทั้งที่โคนต้นลำ ต้น กิ่งและราก โดยจะสังเกตได้จากต้นที่เป็นโรคนี้จะมีใบด้าน ไม่เป็นมันและสีจะค่อยๆ เปลี่ยนเป็นสีเหลืองแล้วร่วงหล่น ต้นที่เป็นโรคจะแสดงอาการเน่าและใบเหี่ยว แผลที่ต้นหรือกิ่งจะเน่าเป็นจุดฉ่ำน้ำ เปลือกจะเน่าเป็นสีน้ำตาลและมีเมือกไหลออกมา ซึ่งจะสังเกตได้ในเวลาเช้าหรือช่วงที่มีอากาศชื้น เมื่อถากเปลือกออกจะเห็นเปลือกด้านในมีสีน้ำตาลแดงหรือน้ำตาลเข้มและถ้าขุดดูรากจะพบว่าที่รากแก้วและรากฝอยถูกทำลายเน่าเป็นสีน้ำตาล ทำ ให้ต้นทุเรียนทรุดโทรมและตายในที่สุด

การป้องกันกำจัด

1. ตัดแต่งกิ่งทุเรียนให้โปร่ง เพื่อให้อากาศถ่ายเทได้สะดวก และเก็บส่วนที่เป็นโรคนำไปเผาทำลาย

2. อย่าให้น้ำขังแฉะบริเวณโคนต้น

3. ถากบริเวณที่เป็นโรคออกเพียงบาง ๆ แล้วใช้สารเคมีจำ พวกเมททาแลคซิลหรือฟอสเอทธิลอะลูมินั่มผสมน้ำทาบริเวณที่ถากออก

2) โรคใบติด

เกิดจากเชื้อรา โรคนี้จะระบาดมากในช่วงฤดูฝน ลักษณะอาการบนใบที่พบจะมีรอยคล้าย ๆ ถูกน้ำร้อนลวก ขอบแผลไม่แน่นอนอาจเริ่มที่ปลายใบ กลางใบ หรือโคนใบ แล้วลุกลามจนเป็น

ทั้งใบ และจะสังเกตเห็นเส้นใยสีขาวนวลแผ่ปกคลุมคล้ายใยแมงมุมแผ่ไปตามผิวใบใบที่ถูกทำลายจะร่วงหล่นไปในที่สุด ถ้าใบที่เป็นโรคไปสัมผัสกับใบที่ปกติไม่ว่าจะเป็นใบที่อยู่ล่าง ๆ หรือใบที่อยู่เหนือกว่า ใบปกตินั้นก็จะเป็นโรคใบติดได้เช่นกัน

การป้องกันกำจัด

1. ไม่ควรปลูกทุเรียนให้ชิดกันเกินไปเพราะจะทำ ให้ทรงพุ่มประสานกัน เกิดเป็นโรคติดต่อกันได้ง่าย

2. ตัดแต่งกิ่งที่เป็นโรคเผาทำลายเสีย ตลอดจนทำ ความสะอาดบริเวณโคนต้นโดยเก็บใบที่เป็นโรคเผาทำลาย

3. ฉีดพ่นด้วยสารเคมี เช่น คาร์เบนดาซิม ไทอะเบนดาโซล หรือคอปเปอร์ออกซีคลอไรด์ พ่นทุก 5-7 วันต่อครั้ง

3) โรคราสีชมพู

เกิดจาก เชื้อราเข้าทำลายกิ่งโดยเฉพาะบริเวณง่ามกิ่งซึ่งมีผลทำ ให้ใบมีสีเหลืองร่วงหล่นไปคล้ายกับอาการกิ่งแห้งและใบร่วงที่เกิดจากโรคโคนเน่าแต่จะสังเกตเห็นเส้นใยของเชื้อรามีลักษณะเป็นขุยสีชมพูปกคลุมบริเวณโคนกิ่งที่มีใบแห้งนั้น และทำ ให้เปลือกของกิ่งทุเรียนปริแตก และล่อนจากเนื้อไม้เมื่อถากเปลือกจะพบว่าเนื้อไม้ภายในมีสีน้ำตาล ถ้าเกิดรอบกิ่งจะทำให้กิ่งทุเรียนแห้งตายในที่สุด

การป้องกันกำจัด

1. ตัดแต่งกิ่งที่เป็นโรคเผาทำลายทิ้งเสีย และตัดแต่งกิ่งทุเรียนให้โปร่งเพื่อให้อากาศถ่ายเทได้สะดวก

2. ฉีดพ่นด้วยสารเคมีจำ พวกคอปเปอร์ออกซิคลอไรด์

4) โรคราแป้ง

เกิดจาก เชื้อราเข้าทำลายผลทุเรียนตั้งแต่เริ่มติดผลอ่อนจนกระทั่งผลแก่ ผิวผลอ่อนจะมีผงสีขาวคล้ายโรยด้วยแป้ง และผลอ่อนก็จะร่วงไปแต่ถ้าเชื้อโรคเข้าทำลาย เมื่อผลโตแล้วจะทำให้ผลแก่มีสีผิวที่ผิดปกติไม่เป็นที่ต้องการของตลาด

การป้องกันกำจัด

1. นำผลทุเรียนที่ร่วงหล่นไปเผาทำลาย

2. ฉีดพ่นด้วยกำมะถันผงละลายน้ำในขณะที่ทุเรียนเริ่มติดผล

5) โรคจุดสนิม

เกิดจาก พืชชั้นต่ำพวกสาหร่ายทำ ความเสียหายให้กับทุเรียนโดยดูดอาหารจากใบ ทำให้ต้นทรุดโทรม อาการของโรคจะพบทั้งที่ใบและกิ่งที่ใบจะปรากฏเป็นจุดหรือดวงสีเทาอ่อนปนเขียว แล้วจะเปลี่ยนเป็นสีน้ำตาลคล้ายสนิมมีลักษณะเป็นขุยคล้ายกำ มะหยี่ ส่วนการทำ ลายที่กิ่งนั้นจะทำให้เปลือกหนาซึ่งนานเข้าจะทำ ให้เปลือกแตก กิ่งแห้งและทรุดโทรมในเวลาต่อมา

การป้องกันกำจัด

1. ตัดกิ่งที่เป็นโรคเผาทำลายทิ้งเสีย
2. ฉีดพ่นด้วยสารเคมีจำพวกคอปเปอร์ออกซิคลอไรด์

6) โรคผลเน่า

เกิดจาก เชื้อราเข้าทำลายผลทุเรียน ตั้งแต่เริ่มติดผลอ่อนจนกระทั่งผลแก่ ผิวผลอ่อนจะมีผงสีขาวคล้ายโรยด้วยแป้ง และผลอ่อนก็จะร่วงไป แต่ถ้าเชื้อโรคเข้าทำลายเมื่อผลโตแล้วจะทำ ให้ผลแก่มีสีผิวที่ผิดปกติ ไม่เป็นที่ต้องการของตลาด การป้องกันกำจัด

1. นำผลทุเรียนที่ร่วงหล่นไปเผาทำลาย
2. ฉีดพ่นด้วยกำมะถันผงละลายน้ำในขณะที่ทุเรียนเริ่มติดผล

- **สรรพคุณทางยา**

เกือบทุกส่วนของทุเรียนนำมาทำเป็นยาสมุนไพรรักษาโรคได้เป็นอย่างดี นำมาอาบแก้ไข้ แก้โรคดีว่าน เป็นยาบำรุงร่างกาย บำรุงเป็นยาแก้น้ำเหลืองเสีย หรือรักษาโรคผิวหนังก็ได้ คุณค่าอาหาร : ทุเรียน 100 กรัม ให้พลังงาน 156 แคลอรี โปรตีน 2.1 กรัม ไขมัน 3.3 กรัม คาร์โบไฮเดรท 29.6 กรัม แคลเซียม 29 มิลลิกรัม ฟอสฟอรัส 34 มิลลิกรัม เหล็ก 1.1 มิลลิกรัม เบต้าแคโรทีน 46 ไมโครกรัม วิตามินบี1 0.16 มิลลิกรัม วิตามินบี2 0.23 มิลลิกรัม ไนอะซิน 2.5 มิลลิกรัม วิตามินซี 35 มิลลิกรัม

(3) มังคุด

- **ลักษณะทางพฤกษศาสตร์ (กิสณะ ตันเจริญ, 2552)**

ชื่อภาษาไทย มังคุด

ชื่อภาษาอังกฤษ Durion

ชื่อวิทยาศาสตร์ *Garcinia mangostana* L.

วงศ์ GUTTIFERAE

มังคุด เป็นไม้ผลยืนต้นขนาดใหญ่ ชอบอากาศร้อนชื้น อุณหภูมิที่เหมาะสมอยู่ในช่วง 25-30 องศาเซลเซียส ความชื้นสัมพัทธ์สูงประมาณ 75-85% ดินควรมีค่าความเป็นกรดเป็นด่าง (pH) ประมาณ 5.5-6.5 และที่สำคัญควรเลือกพื้นที่ปลูกที่มีน้ำเพียงพอตลอดช่วงฤดูแล้ง มังคุดจะให้ผลผลิตประมาณปีที่ 7 หลังปลูก แต่ผลผลิตต่อต้นในระยะแรกจะต่ำ ช่วงที่ให้ผลผลิตดีประมาณ 13 ปีขึ้นไป โดยเฉลี่ย 60 กิโลกรัม/ต้น (น้ำหนักผลเฉลี่ย 80 กรัม/ผล) มังคุดเป็นไม้ผลที่มีระบบรากหาอาหารค่อนข้างลึก ประมาณ 90-120 เซนติเมตร จากผิวดิน ดังนั้นจึงต้องการสภาพแล้งก่อนออกดอกค่อนข้างนาน โดยมีลักษณะดังนี้

ราก : เป็นระบบรากแก้ว มีจำนวนรากแขนงไม่มาก และที่บริเวณปลายรากมีขนรากน้อย

ลำต้น : ลำต้นตรง เปลือกภายนอกมีสีน้ำตาลเข้มจนถึงดำ ภาคในเปลือกประกอบไปด้วยท่อน้ำยางมีลักษณะสีเหลือง

ใบ : ใบมีรูปยาว มีความยาวประมาณ 9-25 ซม. กว้างประมาณ 4.5-10 ซม. ด้านบนมีลักษณะเป็นมันสีเขียวเข้ม ส่วนด้านล่างสีเขียวปนเหลือง แผ่นใบโค้งเล็กน้อย มีตาข้างอยู่บริเวณซอกใบ และมีตายอดอยู่บริเวณซอกใบคู่สุดท้าย

ดอก : เป็นแบบเดี่ยวและบางสภาพอาจเป็นดอกกลุ่ม ซึ่งดอกจะปรากฏที่บริเวณปลายยอดของกิ่งแขนง ที่มีช่อดอกตัวผู้และดอกตัวเมียอยู่ในดอกเดียวกันดอกจัดเป็นดอกสมบูรณ์เพศแต่เกสรตัวผู้จะเป็นหมัน ดอกมังคุดประกอบด้วยกลีบเลี้ยง 4 กลีบ มีกลีบดอกค่อนข้างหนา 4 กลีบดอก เกสรอยู่ที่ฐานรอบๆ ของรังไข่

ผล : เป็นแบบเบอรี่ มีเส้นผ่าศูนย์กลาง 3.4-7.5 ซม. มีเปลือกหนา 6-10 ซม. เนื้อสีขาวขุ่น ลักษณะของผลอ่อนเปลือกนอกจะมีสีเขียวปนเหลือง มียางสีเหลืองอยู่ภายใน ผลหนึ่งๆ จะมีเมล็ดอยู่ประมาณ 1-6 เมล็ด เมล็ดมีความยาวประมาณ 2.5 ซม. และกว้างประมาณ 1.6 ซม.

- **พันธุ์พื้นเมือง**

การเลือกต้นพันธุ์ :เลือกต้นพันธุ์ที่ได้จากการเพาะเมล็ด มีความสมบูรณ์แข็งแรง อายุไม่น้อยกว่า 2 ปี หรือมีความสูง 30 ซม. มีระบบรากสมบูรณ์ ไม่คดงอ

ระยะปลูก :ระยะปลูกระหว่างแถวและต้น 8x8 ม. หรือ10x10 ม.

วิธีปลูก : ขุดหลุมปลูกขนาด 50x50x50 ซม. ระยะห่างระหว่างต้น 10-12 ม. ระยะห่างระหว่างแถว 10-12 ม. นำปุ๋ยคอกหรือปุ๋ยอินทรีย์อื่นๆ ผสมกับดินที่ขุดไว้บนปากหลุม ในอัตราดิน 3 ส่วน/ปุ๋ยคอก 1 ส่วน อาจผสมปุ๋ยเคมีสูตร 15-15-15 ประมาณ 2-3 กำมือ และปุ๋ยร็อคฟอสเฟต 1 กำมือ นำต้นมังคุดวางกลางหลุม กลบดินให้แน่นปักไม้ยึดลำต้น คลุมโคน รดน้ำให้ชุ่ม หลังปลูกควรทำร่มเงาโดยการใช้ทางมะพร้าว หรือใช้วัสดุอื่นๆ ตามความเหมาะสม

- **การดูแลรักษา**

1) การให้น้ำ

ให้น้ำสม่ำเสมอช่วงเจริญทางใบ และงดให้น้ำช่วงปลายฝน ต้นมังคุดที่มีอายุตายอด 9-12 สัปดาห์และผ่านสภาพแล้ง 20-30 วัน เมื่อแสดงอาการใบตก ปลายใบบิด ก้านใบและกิ่งที่ปลายยอดเริ่มเป็นร่อง ให้กระตุ้นการออกดอกโดยการให้น้ำอย่างเต็มที่ให้มากถึง 1,100 – 1,600 ลิตร/ต้น* จากนั้นให้หยุดดูอาการ 7-10 วัน เมื่อพบว่าก้านใบและกิ่งที่ปลายยอดเริ่มเต่งขึ้นก็ให้น้ำ เป็นครั้งที่ 2 ในปริมาณ 1/2 ของครั้งแรก หลังจากนั้น 10-14 วัน ตาดอกจะผลิออกมาให้เห็น และควรมีการจัดการน้ำเพื่อควบคุมให้มีปริมาณดอกเพียงร้อยละ 35-50 ของยอดทั้งหมด เพื่อให้ได้ผลผลิตที่มีคุณภาพ โดยหลังจากมังคุดออกดอกแล้ว 10-15% ของตายอดทั้งหมด ควรให้น้ำปริมาณมาก 220-280 ลิตร/ต้น* ทุกวัน จนพบว่าในยอดที่ยังไม่ออกดอกเริ่มมียอดอ่อนแทนตาดอกจึงค่อยให้น้ำตามปกติ คือ 80-110 ลิตร/ต้น และจะต้องให้น้ำในปริมาณนี้อย่างสม่ำเสมอต่อเนื่องทุกวัน เพื่อให้ผลมังคุดมีพัฒนาการที่ดี

2) การใส่ปุ๋ย

2.1) เพื่อบำรุงต้นหลังการเก็บเกี่ยว

- ปุ๋ยอินทรีย์ 20 - 50 กิโลกรัม/ต้น
- ปุ๋ยเคมีสูตร 15-15-15 หรือ 16-16-16 1 - 3 กิโลกรัม/ต้น

2.2) เพื่อส่งเสริมการออกดอก (ช่วงปลายฝน)

- ปุ๋ยเคมีสูตร 8-24-24 หรือ 9-24-24 2 - 3 กิโลกรัม/ต้น

2.3) เพื่อบำรุงผล (หลังติดผล 3-4 สัปดาห์)

- ปุ๋ยเคมีสูตร 13-13-21 2 - 3 กิโลกรัม/ต้น

3) การปฏิบัติอื่น ๆ

3.1) การเตรียมสภาพต้นให้พร้อม คือ การจัดการให้ต้นมังคุดแตกใบอ่อนในเวลาที่เหมาะสม และพัฒนาเป็นใบแก่ได้พอดีกับช่วงที่มีสภาพแวดล้อมเหมาะสมปกติต้นมังคุดที่ตัดแต่งกิ่งและใส่ปุ๋ยตามคำแนะนำจะแตกใบอ่อนตามเวลาที่เหมาะสม แต่ต้นที่ไว้ผลมากและขาดการบำรุงที่ดีในฤดูที่ผ่านมา แม้จะจัดการต่างๆ แล้วก็มักจะไม่ค่อยแตกใบอ่อนหรือแตกใบอ่อนช้า จึงควรกระตุ้นการแตกใบอ่อนโดยฉีดพ่นปุ๋ยยูเรีย อัตรา 100-200 กรัม/น้ำ 20 ลิตร แต่ถ้าพ่นยูเรียไปแล้วมังคุดยังไม่ยอมแตกใบอ่อนก็ให้ใช้ ไทโอยูเรีย 20-40 กรัม ผสมน้ำตาลเด็กซ์โตรส 600 กรัม ในน้ำ 20 ลิตร (ไทโอยูเรียมีความเป็นพิษต่อพืชสูงจะทำให้ใบแก่ร่วงได้จึงต้องใช้ด้วยความระมัดระวัง) เมื่อมังคุดแตกใบอ่อนแล้วให้ดูแล

รักษาใบอ่อนที่แตกออกมาให้สมบูรณ์ โดยการหมั่นตรวจสอบและป้องกันการระบาดของหนอนกัดกินใบและโรคใบจุดอย่างใกล้ชิด

3.2) การจัดการเพื่อเพิ่มปริมาณผลผลิตที่มีคุณภาพ (มังคุดคุณภาพ หมายถึงผลมังคุดที่มีผิวลายไม่เกิน 5% ของผิวผลและมีน้ำหนักไม่น้อยกว่า 80 กรัม ปราศจากอาการเนื้อแก้วยางไหลภายในผล และจะต้องเป็นผลที่เก็บเกี่ยวถูกวิธี) : ควบคุมปริมาณดอก มังคุดทุกดอกจะเจริญเป็นผลได้โดยไม่ต้องผสมเกสรถ้าปล่อยให้ออกดอกมากเกินไปผลที่ได้มีขนาดเล็กราคาไม่ดีและยังมีผลต่อความสมบูรณ์ของต้นในปีถัดไป นอกจากจะจัดการน้ำตามที่กล่าวแล้ว ในกรณีที่พบว่ามังคุดออกดอกมากเกินไปแล้วให้หว่านปุ๋ยทางดินสูตร 16-16-16 หรือ 15-15-15 ปริมาณ 2 เท่าของปกติควบคู่กับการให้น้ำจะทำให้ผลที่มีอายุ 2-3 สัปดาห์ร่วงได้บางส่วน : ตรวจสอบและป้องกันกำจัดศัตรูเพลี้ยไฟ ไรแดง ไรขาว อย่างใกล้ชิดในช่วงดอกใกล้บาน และติดผลขนาดเล็ก

3.3) การเก็บเกี่ยว เก็บเกี่ยวผลที่แก่พอเหมาะ เมื่อผลเริ่มเป็นระยะสายเลือด คือ ผลที่มีสีเหลืองอ่อนปนเขียว มีจุดประสีชมพูกระจายทั่วผล แนะนำให้เก็บเกี่ยวด้วยตะกร้อผ้าเพื่อป้องกันผลตกกระแทกและรอยขีดข่วนที่ผิว

3.4) การปฏิบัติหลังการเก็บเกี่ยว หลังเก็บเกี่ยวผลมังคุดแล้ว เก็บไว้ในที่ร่ม ทำความสะอาดผล ขูดยางที่เปลือกออก และคัดแยกคุณภาพก่อนจำหน่ายแยกมังคุดเป็น 3 พวก คือมีสีเขียวที่ผิวผล ผิวผลมีสีแดง และผลมีสีม่วงแดงจนถึงม่วงดำ นำมาแยกตามขนาด ทำการคัดผลที่มีตำหนิแยกไว้ต่างหาก ภาชนะที่ใช้ใส่ควรใช้เข่งขนาดเล็กหรือตะกร้าพลาสติก ด้านในควรมีวัสดุ เช่น การดาษ หรือกระสอบปุ๋ยกรุ เพื่อกันช้ำและลดการเสียดสี ไม่ควรบรรจุเกิน 20 กก. ควรวางผลมังคุดไว้ในที่ร่มซึ่งมีอากาศถ่ายเทได้ดี และเมื่อทำการบรรจุเสร็จควรรีบส่งตลาดทันที

- **โรคพืช/ศัตรูพืชของมังคุด**

โรคใบจุด / โรคราสนิม / หนอนกินใบอ่อน / หนอนชอนใบ / เพลี้ยไฟ / ไรแดง

- **การป้องกันกำจัดศัตรูพืช**

1) หนอนกินใบ สุมหญ้าใต้โคนต้นมังคุดเพื่อให้ตัวหนอนมาหลบซ่อนตอนกลางวันและเผาทำลาย ถ้าพบการทำลายมากให้ฉีดพ่นด้วยคาร์บาริล

2) หนอนชอนใบ เก็บใบอ่อนที่มีหนอนเผาทำลาย ถ้าพบการทำลายมากให้ฉีดพ่นด้วย คาร์บาริล

3) โรคใบจุด ในระยะที่แตกใบอ่อนให้ฉีดพ่นด้วยสารเคมี แมนโคเซบคาเบนดาซิม หรือเบนโนมิล ช่วงดอกใกล้บานและติดผลอ่อน

4) เพลี้ยไฟ ฉีดพ่นสารคาร์โบซัลแฟน หรือฟิโปรนิล หรืออิมิลาโคลพริดให้ทั่วทั้งภายในและภายนอกทรงพุ่ม จากนั้น 5-7 วัน ถ้ายิ่งพบว่ามีเพลี้ยไฟให้ฉีดพ่นซ้ำแต่ควรสลับชนิดสารเคมี

5) ไรแดง พ่นด้วยกำมะถันผลหรือสารเคมีโอไมท์ผง

- **ดินที่ปลูกมังคุด (สุรชาติ เพชรแก้, 2542)**

เกษตรกรนิยมปลูกมังคุดมากในพื้นที่ภาคใต้และภาคตะวันออกของประเทศไทย ปัจจุบันเกษตรกรผู้ปลูกมังคุดประสบปัญหาการควบ คุมคุณภาพผลผลิตหลายประการ สาเหตุประการหนึ่งน่าจะมาจากการจัดการธาตุอาหารในดิน ปุ๋ย และต้นมังคุดที่ไม่เหมาะสม สมบัติทางกายภาพของดินส่วนใหญ่เป็นดินร่วนปนทราย และดินมีความสามารถในการระบายน้ำดีจึงมีความเหมาะสมสำหรับเป็นพื้นที่ปลูกมังคุด ซึ่งมังคุดสามารถเจริญเติบโตได้ดีในพื้นที่ที่เป็นดินร่วนปนทรายถึงดินเหนียว ที่มีความสามารถในการระบายน้ำปานกลางถึงดี

- **สรรพคุณทางยา (วันดี กฤษณพันธ์, 2541)**

1) คุณค่าทางโภชนาการ

การบริโภคมังคุด ทำให้เราได้บริโภคกากใยจากเนื้อของมังคุดด้วย ซึ่งจะช่วยในการขับถ่าย และยังได้สารอาหารวิตามินและเกลือแร่ต่างๆ อีกหลายชนิด เช่น น้ำตาล กรดอินทรีย์ แคลเซียม ฟอสฟอรัส และเหล็ก ประโยชน์ของมังคุดมิได้มีอยู่แค่เนื้อในของมังคุดที่เราใช้เป็นอาหารเท่านั้น เปลือกมังคุดก็มีสรรพคุณเป็นยาแก้ท้องเสีย แก้ท้องร่วงเรื้อรัง ถ่ายเป็นมูกเลือด โดยการใช้เปลือกสดหรือเปลือกแห้งฝนกับน้ำรับประทานหรือจะใช้เปลือกแห้งต้มกับน้ำรับประทานก็ได้ผลเช่นเดียวกัน

2) การนำไปใช้ประโยชน์

ปัจจุบันวงการเครื่องสำอางและผลิตภัณฑ์ทำความสะอาดได้ให้ความสนใจนำสารสกัดจากเปลือกมังคุดไปใช้เป็นส่วนผสมในผลิตภัณฑ์ต่างๆ เช่น สบู่เปลือกมังคุด ที่ช่วยดับกลิ่นเต่า ช่วยบรรเทาโรคผิวหนัง รักษาสิวฝ้า ซึ่งใช้ได้ผลดีและเป็นที่นิยมของผู้บริโภค เมื่อได้ลิ้มรสของเนื้อในของมังคุดอย่างอิ่มเอมแล้วก็อย่าได้ทิ้งขว้างเปลือกมังคุดให้เป็นขยะเน่าเหม็นโดยเปล่าประโยชน์เลย เปลือกมังคุดยังมีสรรพคุณในการสมานแผล ช่วยให้แผลหาเร็ว เช่นใช้รักษาบาดแผลผุพอง แผลเน่าเปื่อย แผลเป็นหนอง โดยการใช้เปลือกมังคุดฝนกับน้ำปูนใสทาบริเวณแผล น้ำต้มเปลือกมังคุดแห้งต้มน้ำล้างแผลใช้แทนการด้วยน้ำยาล้างแผลหรือด่างทับทิมได้ด้วย เพราะเปลือกมังคุดนี้มีสารแทนนิน (Tannin) และสารแซนโทน (Xanthone) ที่มีชื่อเรียกเฉพาะชื่อเดียวกับมังคุดว่า สารแมงโกสติน (Mangostin) สารแทนนินมีฤทธิ์สมานแผลช่วยให้แผลหายเร็วขึ้น สารแมงโกสตินมีฤทธิ์ช่วยลดอาการอักเสบ และต้านเชื้อแบคทีเรียที่ทำให้เกิดหนอง สารแซนโทนในเปลือกมังคุดยังมีฤทธิ์ในการยับยั้งเชื้อราที่เป็นสาเหตุของโรคผิวหนังและกลากได้

ภาคผนวก ข
การติดตั้ง RDF API for PHP

การติดตั้ง RDF API for PHP

ในการเชื่อมต่อไฟล์ OWL กับ ภาษา PHP จำเป็นต้องใช้เครื่องมือที่มีชื่อว่า RDF API for PHP หรือเรียกอักษรย่อว่า RAP โดยมีขั้นตอนการติดตั้งดังนี้

1. เข้าสู้เว็บไซต์หลัก http://wifo5-03.informatik.uni-mannheim.de/bizer/rdfapi/

RAP - RDF API for PHP V0.9.6

RAP is a software package for parsing, querying, manipulating, serializing and serving RDF models.

Its features include:

- statement-centric methods for manipulating an RDF model as a set of RDF triples
- resource-centric methods for manipulating an RDF model as a set of resources
- ontology-centric methods for manipulating an RDF model through vocabulary specific methods
- quad- and named graph-centric methods for manipulating RDF datasets
- integrated RDF/XML, N3, N-TRIPLE, TriX, GRDDL, RSS parser
- integrated RDF/XML, N3, N-TRIPLE, TriX serializer
- in-memory or database model storage
- SPARQL query engine supporting all features of the W3C SPARQL Recommendation
- SPARQL client library
- RDQL query engine
- inference engine supporting RDF-Schema reasoning and some OWL entailments
- integrated RDF server providing similar functionality as the Joseki RDF server
- integrated linked data frontend
- graphical user-interface for managing database-backed RDF models
- support for common vocabularies
- drawing graph visualizations

2. เลื่อนลงมายังตำแหน่ง Download and Installation คลิกลิงค์ดาวน์โหลด

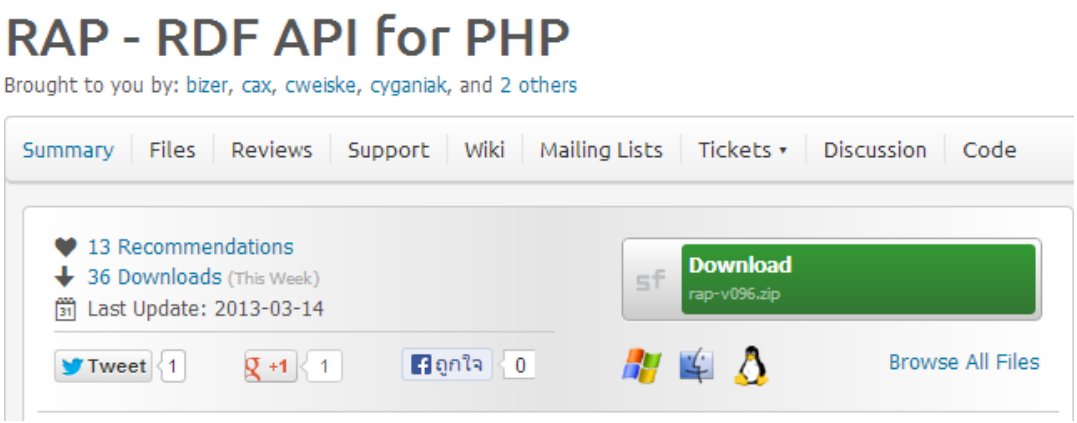
Download and Installation

1. Download the current version from: http://sourceforge.net/projects/rdfapi-php/
2. Unpack the zip file.
3. Include RDF API into your scripts:
   - define("RDFAPI_INCLUDE_DIR", "C:/Apache/htdocs/rdf_api/api/");
   - include(RDFAPI_INCLUDE_DIR . "RDFAPI.php");
4. Change the constant RDFAPI_INCLUDE_DIR to the directory in which you have unpacked the zip file.

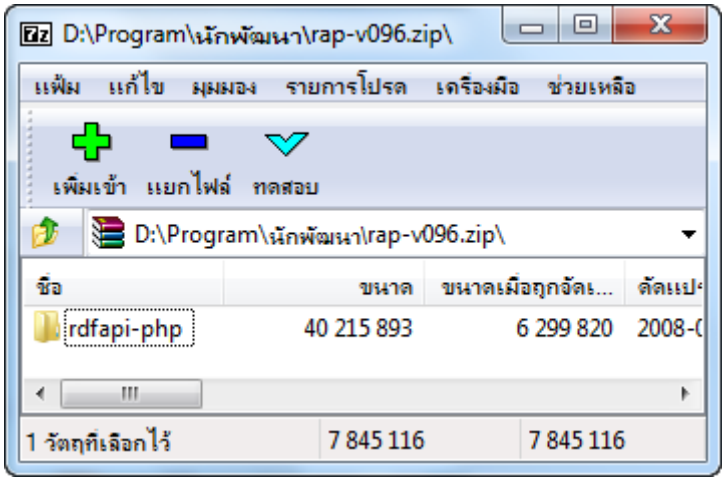
RAP can be used with PHP 4 or PHP 5.

- When used with PHP 4 it requires at least version 4.2
- When used with PHP 5 it requires at least version 5.0.1

3. ดาวโหลดโดยการคลิกปุ่มสีเขียว



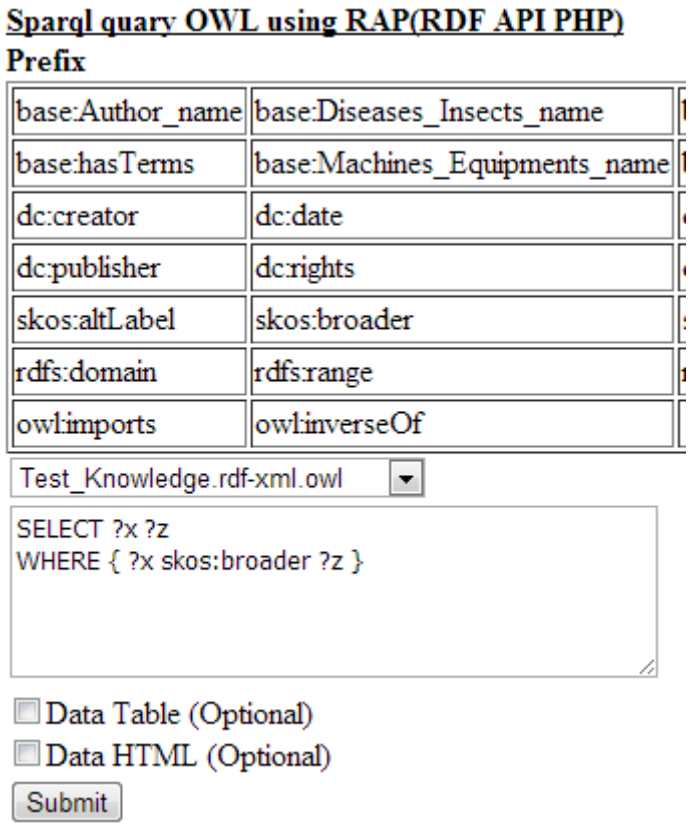
4. แตกไฟล์ zip ไปยัง directory เว็บไซต์ โดยใช้โปรแกรม เช่น 7zip , Zip และ Winrar เป็นต้น



5. สร้าง page เพื่อเรียกใช้ RDF API for PHP สำหรับการสืบค้นข้อมูลภายในไฟล์ OWL ตามคำร้องขอ

```
define("RDFAPI_INCLUDE_DIR", "..\\rdfapi-php\api\\");
include(RDFAPI_INCLUDE_DIR . "RdfAPI.php");

// OWL document Filename
$base=$_POST['select'];

// Create MemModel
$model = new MemModel();

// Load OWL document
$model->load($base);

//Create Prefix string
$baseURI = $model->getBaseURI();
```

ตัวอย่างผลลัพธ์การเข้าสู่ page สำหรับร้องขอคำค้น



ผลลัพธ์การสืบค้นด้วยคำสั่ง SELECT ?x ?z WHERE { ?x skos:broader ?z }

Result :

>> Mangosteen >> Fruits
>> Mangosteen >> Fruit_crops
>> Seedless_varieties >> Varieties
>> Durian >> Fruits
>> Durian >> Fruit_crops
>> Fruit_trees >> Fruit_crops
>> Fruit_trees >> Fruits
>> Drying_mangosteens >> Packaging_materials
>> Introduced_varieties >> Varieties
>> Hybrids >> Varieties
>> Varieties >> Durian
>> Varieties >> Mangosteen
>> Durian_Guan >> Packaging_materials
>> Land_varieties >> Varieties
>> Packaging_materials >> Mangosteen
>> Packaging_materials >> Durian
>> High_yielding_varieties >> Varieties

สำหรับการสืบค้นแบบระบุค่าคงที่จำเป็นต้องเพิ่ม prefix ตัวอย่างเช่น SELECT ?x WHERE { ?x skos:broader base:Fruits } ผลการร้องขอแสดงดังภาพ

Result :
>> Mangosteen
>> Durian
>> Fruit_trees

สำหรับการสืบค้น altinate lable และ prefer labla แบบไม่เจาะจงนั้นสามารถทำได้เช่นเดียวกับการสืบค้นทั่วไป ตัวอย่างเช่น SELECT ?x ?z WHERE { ?x skos:prefLabel ?z } ผลการร้องขอแสดงดังภาพ

Result :
>> Concept_208 >> โครงสร้าง
>> Concept_19 >> x
>> Concept_19 >> xxx
>> Concept_18 >> x
>> Concept_18 >> xxx
>> Mangosteen >> Mangosteen
>> Mangosteen >> มังคุด
>> Seedless_varieties >> Seedless varieties
>> Seedless_varieties >> พันธุ์ไร้เมล็ด
>> □□_ล□□_ม□□_ >> Fruits
>> Tropical_fruits >> Tropical fruits
>> Tropical_fruits >> ผลไม้เขตร้อน
>> Temperate_fruits >> Temperate fruits
>> Temperate_fruits >> ผลไม้เขตหนาว

สำหรับการสืบค้น altinate lable และ prefer labla แบบระบุค่าคงที่จำเป็นต้องเติมคำสั่งตัวอย่างเช่น SELECT ?x WHERE { ?x skos:prefLabel “Seedless varieties”^^xsd:string } ผลการร้องขอแสดงดังภาพ

Result :
>> Seedless_varieties >>

การแสดงข้อมูลไฟล์ OWL ในแบบ Table สามารถทำให้มองเห็นข้อมูลได้อย่างชัดเจน ซึ่งใน RAP ได้เตรียมความสามารถนี้ไว้ให้แล้ว โดยการเรียกคำสั่ง $model->writeAsHtmlTable(); ผ่านตัวแปร MemModel

Table :

Base URI: http://www.owl-ontologies.com/Ontology1348489755.owl#	
Prefix:	**Namespace:**
xsp	http://www.owl-ontologies.com/2005/08/07/xsp.owl#
swrlb	http://www.w3.org/2003/11/swrlb#
swrl	http://www.w3.org/2003/11/swrl#
protege	http://protege.stanford.edu/plugins/owl/protege#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
xsd	http://www.w3.org/2001/XMLSchema#
rdfs	http://www.w3.org/2000/01/rdf-schema#
owl	http://www.w3.org/2002/07/owl#
skos	http://www.w3.org/2004/02/skos/core#
dc	http://purl.org/dc/elements/1.1/
No.	**Subject**
1.	Resource: http://www.owl-ontologies.com/Ontology1348489755.owl
2.	Resource: http://www.owl-ontologies.com/Ontology1348489755.owl

การแสดงข้อมูลไฟล์ OWL ในแบบ HTML สามารถทำให้มองเห็นข้อมูลได้เช่นเดียวกับที่ ซึ่งใน RAP ได้เตรียมความสามารถนี้ไว้ให้แล้ว โดยการเรียกคำสั่ง $model-> writeAsHtml(); ผ่านตัวแปร MemModel

HTML :

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Generated by RdfSerializer.php from RDF RAP.
# http://www.wiwiss.fu-berlin.de/suhl/bizer/rdfapi/index.html !-->

<rdf:RDF
  xml:base="http://www.owl-ontologies.com/Ontology1348489755.owl#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:vcard="http://www.w3.org/2001/vcard-rdf/3.0#"
  xmlns:ns1="http://www.owl-ontologies.com/Ontology1348489755.owl#"

<rdf:Description rdf:ID="□□_ัดลม□□_ร□□_ดั□□_สู□□_">
  <ns1:Machines_Equipments_name rdf:datatype="http://www.w3.org/200
แรงดันสูง</ns1:Machines_Equipments_name>
</rdf:Description>
```

ภาคผนวก ค
การหาคุณภาพและตัวอย่างแบบประเมินความสามารถของระบบ

แบบประเมินการทำงานของระบบการค้นหาเชิงความหมายสำหรับไม้ผลเศรษฐกิจไทย
การวิจัยเรื่อง การเข้าถึงความรู้ทางการเกษตรด้วยเทคโนโลยีเว็บเชิงความหมาย

.....

คำชี้แจง

1. แบบสอบถามชุดนี้ แบ่งออกเป็น 2 ตอน ประกอบด้วย

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม จำนวน 2 ข้อ

ตอนที่ 2 ความคิดเห็นเกี่ยวกับการทำงานของระบบสืบค้นเชิงความหมายสำรับไม้ผลเศรษฐกิจไทย จำนวน 5 ด้าน รวม 20 ข้อ

2. การตอบแบบสอบถามนี้ คำตอบของท่านมีความสำคัญอย่างยิ่งต่อการวิเคราะห์ข้อมูลที่ใช้ในการวิจัย กรุณาตอบข้อมูลให้ครบถ้วนทุกข้อ ทุกตอน เพื่อให้การวิจัยในครั้งนี้มีความเที่ยงตรงและเกิดประโยชน์อย่างแท้จริง

3. ข้อมูลที่ท่านตอบแบบสอบถามนี้จะเก็บไว้เป็นความลับ การนำเสนอผลการวิจัยจะเสนอในภาพรวมเท่านั้น

ขอขอบพระคุณที่ให้ข้อมูลไว้ ณ โอกาสนี้

(ดร.ศุภกฤษฎ์ นิวัฒนากูล)
สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม
มหาวิทยาลัยเทคโนโลยีสุรนารี

ตอนที่ 1 ข้อมูลผู้ตอบแบบสอบถาม

1.1 เพศ ☐ 1. ชาย ☐ 2. หญิง

1.2 ระดับการศึกษาสูงสุด

☐ 1. ต่ำกว่าปริญญาตรี ☐ 2. ปริญญาตรีหรือเทียบเท่า

☐ 3. ปริญญาโท ☐ 4. ปริญญาเอก

ตอนที่ 2 ความคิดเห็นเกี่ยวกับการทำงานของระบบ

รายการคำถาม	**ระดับความคิดเห็น**				
	มากที่สุด	**มาก**	**ปานกลาง**	**น้อย**	**น้อยที่สุด**
1. ด้านการประเมินความสามารถของระบบ (Functional requirement test)					
1.1 ความสามารถของส่วนนำเข้าข้อคำถามตรงตามความต้องการของท่านในระดับใด					
1.2 ผลลัพธ์การค้นหาเป็นข้อมูลปัจจุบันและเป็นประโยชน์ต่อผู้ใช้					
1.3 ผลลัพธ์การค้นหาตรงตามความต้องการของผู้ใช้					
1.4 ระบบสามารถเข้าถึงเอกสารความรู้ด้านไม้ผลเศรษฐกิจ					
2. ด้านการประเมินความถูกต้องของระบบ (Functional test)					
2.1 ความถูกต้องของข้อมูลสำหรับการค้นหาแบบหนึ่งคำค้น					
2.2 ความถูกต้องของข้อมูลสำหรับการค้นหาแบบสองคำค้น					
2.3 ความถูกต้องของข้อมูลสำหรับการค้นหาแบบสามคำค้น					
2.4 ความถูกต้องของข้อมูลสำหรับการค้นหาแบบมากกว่าสามคำค้น					

รายการคำถาม	ระดับความคิดเห็น				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
3. ด้านการประเมินความง่ายในการใช้งาน (Usability test)					
3.1 ขนาด สี รูปแบบตัวอักษรมีความเหมาะสม					
3.2 การออกแบบภาพกราฟิก มีความชัดเจนเหมาะสม					
3.3 การจัดวางองค์ประกอบของระบบดูเหมาะสม					
3.4 การเชื่อมโยงไปยังเอกสารความรู้ที่ค้นพบมีความเหมาะสม					
3.5 รูปแบบของการระบุคำค้นสะดวกต่อการใช้งาน					
3.6 ความเหมาะสมในการปฏิสัมพันธ์โต้ตอบกับผู้ใช้					
4. ด้านการประเมินประสิทธิภาพของระบบ (Performance test)					
4.1 สามารถค้นหาข้อมูลได้อย่างรวดเร็ว					
4.2 ระบบมีความเสถียรภาพในการเข้าใช้งาน					
4.3 ระบบมีความแม่นยำในการค้นหา					
5. ด้านการประเมินด้านการรักษาความปลอดภัย (Security test)					
5.1 การกำหนดรหัสผู้ใช้ และรหัสผ่านในการตรวจสอบผู้ดูแลระบบ					
5.2 การตรวจสอบสิทธิ์ก่อนการใช้งานของผู้ดูแลระบบ					
5.3 การป้องกันการล้มเหลวของข้อมูลในระบบ					

ความคิดเห็นอื่น ๆ

.....

.....

.....

.....

.....

(ลงชื่อ).....ผู้ประเมิน

(.....)

ความคิดเห็นของผู้เชี่ยวชาญ คนที่.....

ตารางที่ ค.1 แสดงค่าดัชนีความสอดคล้อง (IOC) แบบประเมินการทำงานของระบบของผู้เชี่ยวชาญที่มีต่อระบบสืบค้นไม้ผลเศรษฐกิจ

ข้อคำถาม	คะแนนความคิดเห็นของผู้เชี่ยวชาญ			รวม	ค่าเฉลี่ย	สรุปผล
	คนที่ 1	คนที่ 2	คนที่ 3			
1.1	1	1	1	3	1.00	ผ่านเกณฑ์
1.2	1	1	1	3	1.00	ผ่านเกณฑ์
1.3	1	1	1	3	1.00	ผ่านเกณฑ์
1.4	1	1	1	3	1.00	ผ่านเกณฑ์
2.1	1	1	1	3	1.00	ผ่านเกณฑ์
2.2	1	1	1	3	1.00	ผ่านเกณฑ์
2.3	1	1	1	3	1.00	ผ่านเกณฑ์
2.4	1	1	1	3	1.00	ผ่านเกณฑ์
3.1	1	1	1	3	1.00	ผ่านเกณฑ์
3.2	1	1	1	3	1.00	ผ่านเกณฑ์
3.3	1	1	1	3	1.00	ผ่านเกณฑ์
3.4	1	1	1	3	1.00	ผ่านเกณฑ์
3.5	1	1	1	3	1.00	ผ่านเกณฑ์
3.6	1	1	1	3	1.00	ผ่านเกณฑ์
4.1	1	1	1	3	1.00	ผ่านเกณฑ์
4.2	1	1	1	3	1.00	ผ่านเกณฑ์
4.3	1	1	1	3	1.00	ผ่านเกณฑ์
5.1	0	1	1	2	0.67	ผ่านเกณฑ์
5.2	0	1	1	2	0.67	ผ่านเกณฑ์
5.3	1	1	1	3	1.00	ผ่านเกณฑ์

ประวัติผู้วิจัย

ชื่อ นายศุภกฤษฎิ์ นิวัฒนากูล
Mr.Suphakit Niwattanakul

ตำแหน่ง
อาจารย์ประจำสาขาวิชาเทคโนโลยีสารสนเทศ
สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี

ที่อยู่
สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม
มหาวิทยาลัยเทคโนโลยีสุรนารี ต.สุรนารี อ.เมือง จ.นครราชสีมา 30000
โทรศัพท์ 044-224560 โทรสาร 044-224205
E-Mail suphakit@ccs.sut.ac.th

ประวัติการศึกษา

- Ph.D. (Computer Science) University of La Rochelle, France (2008)
- วท.ม. (เทคโนโลยีการจัดการระบบสารสนเทศ) มหาวิทยาลัยมหิดล (2542)
- ทษ.บ. (บริหารธุรกิจการเกษตร) มหาวิทยาลัยแม่โจ้ เชียงใหม่ (2533)

งานวิจัย

- การเข้าถึงความรู้ด้วยตัวแบบออนโทโลยี (Access to Knowledge based on an Ontology Model), 2551, วิทยานิพนธ์ระดับปริญญาเอก : มหาวิทยาลัยลาโรแชล ประเทศฝรั่งเศส
- การพัฒนาต้นแบบสำหรับการสร้างแบบทดสอบผ่านระบบเครือข่ายอินเตอร์เน็ต (Web Based Testing Model Via Internet), 2546 (ได้รับเงินอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี)
- ระบบการสอนเสริมที่ชาญฉลาดสำหรับการเรียนรู้โครงสร้างข้อมูล (Intelligence Tutoring System for Data Structure Learning), 2542, วิทยานิพนธ์ระดับปริญญาโท : มหาวิทยาลัยมหิดล