

บทความวิจัย (Research Article)**การออกแบบออนโทโลยีข้าวไทย****Ontology design of Thai Rice**

ขนิษฐา กุลนาวิน^{1*}, ธนิตา วงศ์กาฬสินธุ์² และ นพรัตน์ ยติกร³
Kanitta Kulnawin^{1*}, Thanita Wongkalasin² and Nopparat Yatigon³

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่ออธิบายกระบวนการออกแบบออนโทโลยีข้าวไทยเพื่อเป็นฐานองค์ความรู้ช่วยค้นคำจากคำสำคัญที่เกี่ยวกับข้าวไทย โดยมีกระบวนการออกแบบ 2 ส่วนคือ กระบวนการออกแบบออนโทโลยี และการประเมินการวัดประสิทธิภาพการค้นคืน ขั้นตอนของกระบวนการพัฒนาออนโทโลยี แบ่งออกเป็น 4 ขั้นตอนคือ 1 การนิยามขอบเขต 2 การออกแบบคลาสและคุณสมบัติ 3 ข้อบังคับในการกำหนดข้อจำกัด 4 การสร้างสมาชิกของคลาส องค์ความรู้เกี่ยวกับข้าวไทยจากฐานข้อมูลขนาดใหญ่จะช่วยเพิ่มประสิทธิภาพในการสืบค้น โดยมีการเชื่อมโยงความสัมพันธ์ข้อมูลที่เกี่ยวข้องกันไว้ในกลุ่มเดียวกันโดยการประยุกต์ใช้เทคนิคการสืบค้นข้อมูล การประเมินวัดประสิทธิภาพของการสืบค้นสามารถคำนวณค่าเฉลี่ยของความแม่นยำได้ 76.29% ผลที่ได้จากการออกแบบออนโทโลยีข้าวไทยสามารถนำไปพัฒนาในการใช้งานตามวัตถุประสงค์ได้ต่อไปในอนาคต

คำสำคัญ: ข้าวไทย, ออนโทโลยี, ความแม่นยำ

Abstract

Article is intended to describe the process of ontology design of Thai rice as a knowledge base for the internet strategies searching of Thai rice keywords. There are two parts of the design process: ontology design phase and ontology evaluation phase. The ontology design phase consists of four steps: design the scope of ontology, class design and features, regulation to set limits and efficiency of searching knowledge of Thai rice. The design will enhance creating class member Thai rice from large database. This method can meet the information needs of user by relating the relevant information into the same application. Ontology evaluation phase shows the average of Precision at 76.29%. The result of the design can be applied to develop a particular purpose in the future.

Key words: Thai rice, ontology, precision

¹ โปรแกรมวิชาวิทยาการสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา จังหวัดนครราชสีมา 30000

¹ Informatics Program, Faculty of Sciences and Technology, Nakhon Ratchasima Rajabhat University, Nakhon Ratchasima Province 30000

E-mail : nonone007@yahoo.com

Received: 28 June 2013; Accepted: 30 March 2014

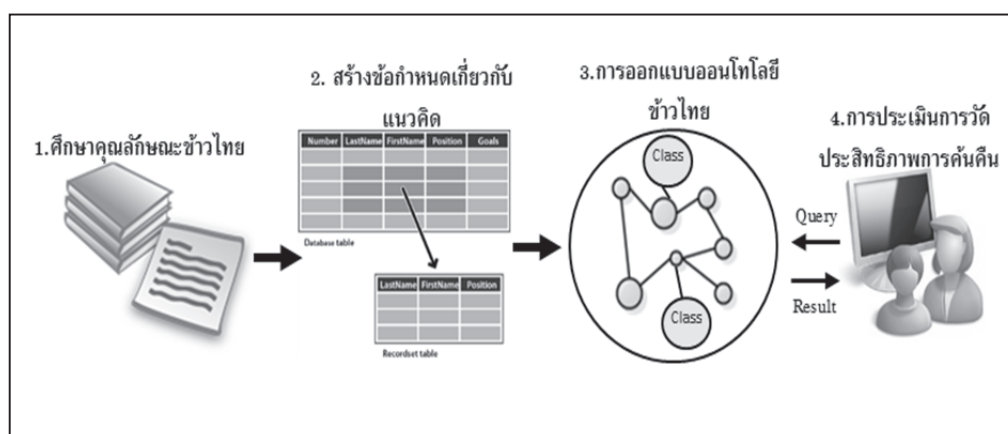
บทนำ

จากข้อมูลทางประวัติศาสตร์พบว่าคนไทยนิยมบริโภคข้าวตั้งแต่สมัยก่อนประวัติศาสตร์จนถึงสมัยรัตนโกสินทร์ ข้าวพันธุ์พื้นเมืองที่เป็นที่นิยมบริโภคไปทั่วโลก คือ ข้าวหอมมะลิ ซึ่งมีลักษณะเมล็ดข้าวใสเหมือนดอกมะลิ มีกลิ่นหอมเหมือนใบเตย ข้าวมีความเกี่ยวพันกับความเป็นชาติไทย คนไทย รวมทั้งวัฒนธรรม ประเพณี วิถีชีวิตคน เศรษฐกิจ กรรมวิธีการปลูกข้าวของไทยเป็นสิ่งที่บรรพบุรุษได้ทดลองคัดเลือก และผสมผสานไว้อย่างเหมาะสม สืบทอดกันมาจนนับพันปี โดยข้าวไทยก็มีความแตกต่างกันไปตามแต่ละภูมิภาค ตามวิธีการปลูก ตามการการตอบสนองต่อช่วงแสง และลักษณะของลำต้น ทำให้ข้าวไทยมีความหลากหลาย สัมพันธ์กับทรัพยากรสภาพภูมิประเทศ ซึ่งเกิดจากภูมิปัญญาของเกษตรกรไทย [1,3] เพื่อก่อให้เกิดการเชื่อมโยงขององค์ความรู้ข้อมูลเรื่องข้าวไทยซึ่งเป็นภูมิปัญญาท้องถิ่นที่อยู่คู่คนไทยมาช้านาน โดยสามารถสืบสานองค์ความรู้ที่เกี่ยวกับคุณลักษณะของข้าวไทยให้อยู่ในลักษณะของฐานองค์ความรู้และสร้างความสัมพันธ์ระหว่างข้อมูลเพื่อให้ข้อมูลอยู่ในลักษณะที่สืบค้นและเข้าถึงได้ง่าย

ในการทดลองนี้จึงได้ใช้เทคโนโลยีออนโทโลยี ซึ่งเป็นเทคโนโลยีทางการพัฒนาภาษาเชิงความหมายสมัยใหม่ ซึ่งเป็นภาษาที่เครื่องคอมพิวเตอร์สามารถตีความหมายและทำตามคำสั่งได้ โดยอาศัยคุณลักษณะของออนโทโลยี (ontology) ที่ต้องมีการนิยามรูปแบบ (model) ภายในขอบเขตขององค์ความรู้เพื่ออธิบายสิ่งที่เราสนใจ (domain) ให้ได้ใจความและถูกต้องมากที่สุด

วัสดุอุปกรณ์และวิธีการศึกษา

การออกแบบออนโทโลยีข้าวไทยเป็นการจับคู่ระหว่างข้อมูลที่มีอยู่ในระบบฐานข้อมูลกับแนวคิดของออนโทโลยี เพื่อการพัฒนารูปแบบของข้อมูลให้สามารถนำไปใช้ประโยชน์ในระบบการสืบค้นเชิงความหมายได้อย่างมีประสิทธิภาพมากยิ่งขึ้นต่อไป [10] ขั้นตอนของการออกแบบออนโทโลยีข้าวไทย เริ่มจาก 1) ศึกษาคุณลักษณะของข้าวไทย 2) สร้างข้อกำหนดเกี่ยวกับแนวคิดของข้าวไทย 3) ออกแบบออนโทโลยีข้าวไทย และ 4) การประเมินการวัดประสิทธิภาพการค้นคืน (ภาพที่ 1)



ภาพที่ 1 แสดงภาพรวมของการออกแบบออนโทโลยีข้าวไทย

ออนโทโลยีเป็นเทคโนโลยีของเว็บ 3.0 หรือเว็บเชิงความหมาย (semantic web) ซึ่งเข้ามาช่วยในเรื่องการจัดเก็บข้อมูล ที่มีการนำเสนอเนื้อหาแบบมีโครงสร้าง มีขอบเขตเนื้อหาในด้านใดด้านหนึ่งที่สนใจ อีกทั้งยังสามารถวิเคราะห์ จำแนก หรือจัดแบ่งหมวดหมู่ของข้อมูลตามความสัมพันธ์ของชาวไทย [10] ในการออกแบบเริ่มจากศึกษาคุณสมบัติของชาวไทย ซึ่งเป็นพืชจำพวกธัญพืชที่สามารถกินเมล็ดได้ เป็นพืชใบเลี้ยงเดี่ยวเช่นเดียวกับหญ้าที่สามารถกินเมล็ดได้ ต้นข้าวมีลักษณะภายนอกบางอย่าง เช่น ใบ กาบใบ ลำต้น และรากคล้ายต้นหญ้า [5] ในประเทศไทยมีข้าวอยู่หลากหลายสายพันธุ์รวมทั้งข้าวพันธุ์พื้นเมืองและข้าวหอมมะลิที่นิยมไปทั่วโลก คุณลักษณะของชาวไทยได้ทำการรวบรวมข้อมูลจากหนังสือ ฐานข้อมูลเชื้อพันธุ์พืช : ข้าว [6] ข้อมูลทั้งหมดนำมาสร้างเป็นฐานความรู้สำหรับประมวลผล ซึ่งประกอบด้วยคลังคำศัพท์ที่เป็นฐานความรู้พื้นฐานในการประมวลผล จากข้อมูลทั้งหมด 5,329 เรคคอร์ด (record) และ 55 คุณสมบัติ (concept)

ขั้นถัดไปได้ใช้หลักการของเทคโนโลยีออนโทโลยีเพื่อกำหนดความสัมพันธ์ ข้อกำหนดเกี่ยวกับแนวคิด (concepts) ที่บรรยายองค์ความรู้และแสดงข้อมูลที่เกี่ยวข้องกับคุณลักษณะของชาวไทยภายใต้โดเมนและความสัมพันธ์ระหว่างสิ่งเหล่านั้นอย่างมีขอบเขต ตามแนวความคิดของออนโทโลยีซึ่งสามารถแสดงออกมาในรูปของระบบสัญลักษณ์ (notation) ยกตัวอย่างเช่น คลาส (class) อินสแตนซ์ (instance) ความสัมพันธ์ (relationship) คุณสมบัติของคลาส (properties/attribute) และกฎ (rule) โดยใช้ภาษาสำหรับแสดงความรู้ (knowledge representation language) ซึ่งมีความชัดเจนและเที่ยงตรง มากกว่าการอธิบายโดยใช้ภาษาธรรมชาติ (natural language) ที่ใช้คำศัพท์มาเชื่อมต่อกันเป็นประโยคเพื่อบรรยายถึงคุณลักษณะของชาวไทยในแง่มุมต่างๆ ทั้งนี้การใช้ระบบสัญลักษณ์จะช่วยสื่อความหมาย เชื่อมโยงข้อมูลเข้าด้วยกันด้วยโปรแกรมโปรทีเจ 3.4.8 (protégé 3.4.8) เป็นเครื่องมือที่สนับสนุนการพัฒนาออนโทโลยี เพื่อสามารถถ่ายทอดและจัดเก็บองค์ความรู้ในรูปแบบอยู่ในรูปแบบที่เข้าใจและสามารถ

นำไปใช้ประโยชน์ได้ นอกจากนี้ออนโทโลียังสร้างความสัมพันธ์โครงสร้างความรู้ให้อยู่ในรูปแบบลำดับชั้นเชิงวัตถุ (hierarchical data structure) เพื่ออธิบายขอบเขตขององค์ความรู้ชาวไทย วิธีการแก้ปัญหาดังกล่าวคือเปรียบเทียบรายละเอียดของคำที่ต่างกันแล้วแปลงให้สามารถใช้ข้อมูลร่วมกันได้ [13, 17, 18]

การบรรยายขอบเขตความรู้ของชาวไทยเรียกว่า โดเมน (domain) เช่น ชื่อพันธุ์ (ไทย), ชื่อพันธุ์ (อังกฤษ), ความยาวของแผ่นใบ (ซม.), ความกว้างของแผ่นใบ (ซม.), ขนบนแผ่นใบ และปริมาณโปรตีนในข้าวกล้องเป็นต้น (ตารางที่ 1) โดยการอธิบายความรู้ตามขอบเขตที่สนใจ คือการระบุหน้าที่หรือลักษณะที่เป็นส่วนสำคัญของข้อมูลนั้นๆ หากมีการบรรยายขอบเขตความรู้ที่สนใจแล้วยังต้องกำหนดกฎ (rule) ที่เกี่ยวข้องกับโดเมนนั้น หรืออาจนำเสนอกฎเกณฑ์ในการบรรยายหรืออธิบายความหมายของสิ่งที่สนใจได้ด้วยกฎเกณฑ์ทางคณิตศาสตร์ [1, 2, 4] สำหรับการกำหนดองค์ความรู้สำหรับออนโทโลยีต้องมีเหตุผลประกอบ [10 - 12, 14] ดังต่อไปนี้

1. กฎการเป็นสมาชิกของคลาส : ถ้ากำหนดให้ x เป็นวัตถุของคลาส C และคลาส C เป็นคลาสลูก (Sub class) ของคลาส D ดังนั้นสามารถอนุมานได้ว่า x เป็นวัตถุของคลาส D ด้วย
2. กฎการเท่ากันของคลาส : ถ้าคลาส A เทียบเท่ากับคลาส B และคลาส B เทียบเท่ากับคลาส C แล้วคลาส A จะเทียบเท่ากับคลาส C ด้วย
3. กฎการสอดคล้อง : หากประกาศตัวแปร x ให้เป็นวัตถุของคลาส A และ A เป็นคลาสลูกของคลาส $B \cap C$ ซึ่ง A เป็นคลาสลูกของคลาส D และคลาส B กับคลาส D ไม่สัมพันธ์กัน ดังนั้น ไม่สอดคล้องเพราะคลาส A ควรจะไม่มีค่าใดๆ แต่คลาส A มีวัตถุ x อยู่
4. การจัดแบ่งประเภท : ถ้าประกาศค่าคุณสมบัติ (property-value) ครบคู่สมาชิกของคลาส A ตามเงื่อนไขที่แน่นอนแล้ว หากตัวแปร x กำหนดตามเงื่อนไข ดังนั้นสามารถสรุปได้ว่า x ต้องเป็นวัตถุของคลาส A ด้วยขั้นตอนของการออกแบบออนโทโลยีชาวไทย [7 - 11] มีดังนี้

ตารางที่ 1 แสดงตัวอย่างของข้อมูลที่ใช้การทดลองสร้างฐานองค์ความรู้ข้าวไทย

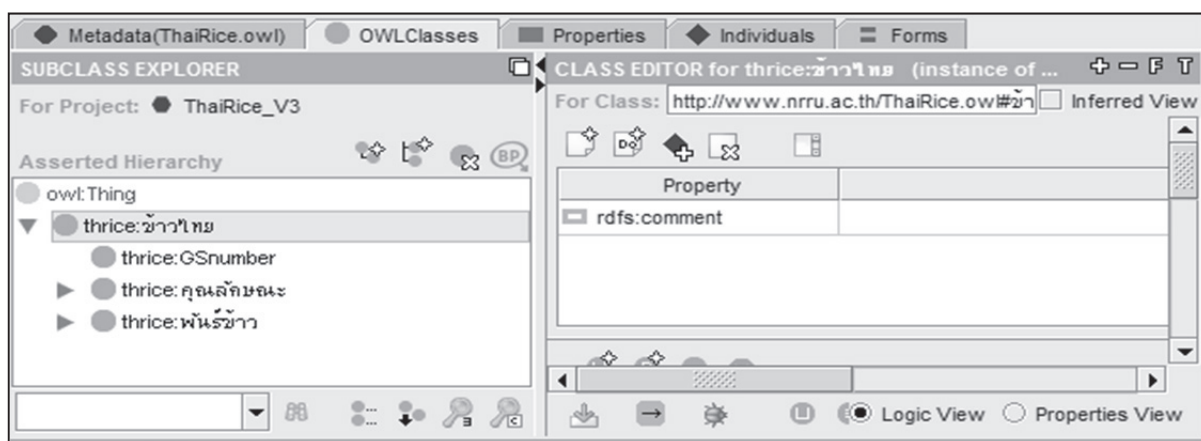
จีเอสเอ็มเบอร์ G.S.number	ชื่อพันธุ์ (ไทย)	ชื่อพันธุ์ (อังกฤษ)	ความยาว ของแผ่น ใบ(ซม.)	ความกว้าง ของแผ่นใบ (ซม.)	ขนบน แผ่นใบ	สีแผ่น ของใบ	...	ปริมาณ โปรตีนใน ข้าวกล้อง
00009	เจ้าหอม 61	JAO HAWM 61	46.00	1.04	มีบ้าง	เขียว เข้ม	...	-
00013	คอเหลือง 106	DAW LEUANG 106	55.40	1.18	มีบ้าง	เขียว เข้ม	...	-
00015	คอเหลือง 150	DAW LEUANG 150	-	-	เกลี้ยง	เขียว	...	-
00017	เจ๊กกระโดด 85-4-34	JEK GRA DOD 85-4-34	54.60	1.54	มีบ้าง	เขียวจาง	...	-
00021	เหลืองใหญ่ 49-2-119	LEUNG YAI 49-2-119	52.60	1.34	มีบ้าง	เขียว	...	-
00022	เหลืองใหญ่ 28-72	LEUNG YAI 28-72	40.20	1.04	มีบ้าง	เขียว	...	-
00024	เหมยหนอง 62 M	MUEY NAWNG 62 M	44.20	1.04	มีบ้าง	ม่วงผสม เขียว	...	-
00025	ลายหล้า	LAI LAH	48.00	1.30	มีบ้าง	เขียว	...	-

ขั้นตอนที่ 1 การนิยามขอบเขต เมื่อทราบข้อมูลที่สนใจ ทำการนิยามขอบเขตของโดเมนให้ชัดเจน อาทิ ลักษณะของงาน รูปแบบคำถามที่ออนโทโลยีต้องสามารถตอบได้ โดยการสร้างรายการคำถามเพื่อช่วยในการทำความเข้าใจลักษณะของออนโทโลยีที่จะออกแบบ

ขั้นตอนที่ 2 การออกแบบคลาสและคุณสมบัติ เป็นขั้นตอนสร้างฐานความรู้ด้วยโปรแกรมโปรตีเจ 3.3.8 ซึ่งเป็นเครื่องมือในการสร้างออนโทโลยีด้วยภาษา RDF ซึ่งเป็นภาษามาตรฐานที่ใช้อธิบายลักษณะและความสัมพันธ์ของข้อมูลคล้ายกับแผนภาพ

อีอาร์ และ ภาษา OWL เป็นภาษาอธิบายข้อมูลในเชิง Ontology โดยพิจารณาจากความสัมพันธ์ระหว่างสิ่งต่างๆ ของความหมาย ลักษณะของโปรแกรมโปรตีเจ 3.3.8 เป็นแบบเปิดรหัสต้นฉบับ และไม่เสียค่าใช้จ่ายในการใช้งาน ในการออกแบบออนโทโลยีข้าวไทยให้มีเนื้อหาครอบคลุมข้อมูล [1, 21, 22] การออกแบบโครงสร้างฐานองค์ความรู้ออนโทโลยี เริ่มต้นดังนี้

กำหนดขอบเขตของออนโทโลยีโดยเริ่มสร้างคลาสของออนโทโลยีองค์ความรู้ข้าวไทย ประกอบไปด้วยคลาสจำนวน 3 คลาส ดังนี้ 1) คุณลักษณะ 2) พันธุ์ข้าว 3) GSNumber (ภาพที่ 2)



ภาพที่ 2 แสดงภาพรวมของคลาสออนโทโลยีองค์ความรู้ข้าวไทย

ขั้นตอนที่ 3 ขอบบังคับในการกำหนดข้อจำกัด เป็นการกำหนดคำศัพท์ที่จะใช้สำหรับนิยามชื่อคลาสและคุณสมบัติของคลาส โดยการกำหนดคอนเซ็ปต์ในออนโทโลยี ประกอบด้วย 2 ส่วนดังนี้ 1) การกำหนดเรนจ์ของคุณสมบัติคลาสที่ไม่ซับซ้อน 2) พยายามกำหนดให้คุณสมบัติของซับคลาส (sub class) เพื่อลดความซับซ้อนของคุณสมบัติคลาส แต่ละคลาสดังกล่าวประกอบไปด้วย ซับคลาสดังนี้

1. ภายใต้ซับคลาสคุณลักษณะ มี 3 ซับคลาส ดังนี้ 1) ต้น 2) เมล็ด 3) สี และ 4) ใบ
2. ภายใต้ซับคลาสเมล็ด มี 2 ซับคลาสน้อย ดังนี้ 1) เมล็ดมีเปลือก และ 2) เมล็ดข้าว
3. ภายใต้ซับพันธุ์ข้าว มี 3 ซับคลาส ดังนี้ 1) ชื่อพันธุ์อังกฤษ และ 2) ชื่อพันธุ์ไทย

ดังแสดงด้วยภาพที่ 3 ที่แสดงตัวอย่างของแบบจำลองออนโทโลยีองค์ความรู้ข้าวไทยด้วยโปรแกรมโปรทีเจ 3.3.8

ขั้นตอนที่ 4 การสร้างสมาชิกของคลาส เป็นการสร้างสมาชิกให้แก่คลาสโดยการกำหนดค่าให้แก่คุณสมบัติคลาสที่เหมาะสมเพื่อให้เกิดการเชื่อมโยงระหว่างข้อมูล ดังแสดงข้อมูลในตารางที่ 2

เมื่อข้อมูลมีความสัมพันธ์กันดังแสดงด้วยภาพที่ 3 เวลาทำการสืบค้นข้อมูลก็จะสามารถทำประมวลผลการสืบค้นข้อมูลโดยกระบวนการค้นหาเชิงความหมาย และแสดงผลลัพธ์ที่ได้นั้นให้กับผู้ใช้งาน ตัวอย่างเช่น ต้องการสืบค้นข้อมูลที่มีคำว่า “ข้าวเจ้า” ไม่ว่าจะเป็นส่วนหนึ่งของประโยคหรือคำขึ้นต้นประโยค ระบบจะสามารถสืบค้นข้อมูลที่มีความสัมพันธ์กับ “ข้าวเจ้า” และสามารถแสดงผลได้อย่างเป็นลำดับขั้น

Property	Value
rdfs:comment	
	thrice:ข้าวไทย
	thrice:คุณลักษณะ
	thrice:พันธุ์ข้าว

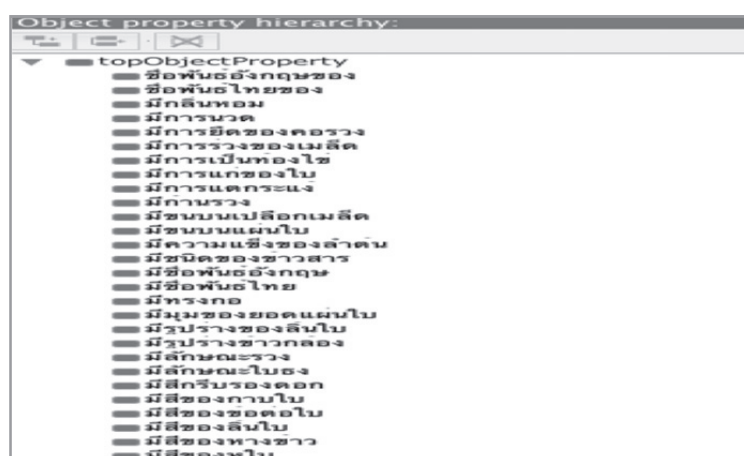
ภาพที่ 3 ตัวอย่างของแบบจำลองออนโทโลยีองค์ความรู้ข้าวไทยด้วยโปรแกรมโปรทีเจ 3.3.8

ตารางที่ 2 ตัวอย่างรายละเอียดของคุณสมบัติของออบเจ็ค (object property) ในออนโทโลยีองค์ความรู้ข้าวไทย

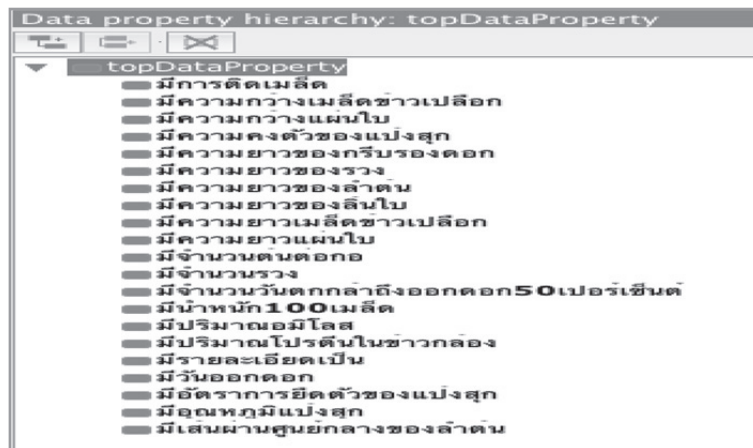
คุณสมบัติ ที่	คุณสมบัติของออบเจ็ค (object property)	โดเมน (domain)	เรนจ์ (range)	อินเวอร์ส (inverse)	ทรานสิทีฟ (transitive)
1.	ชื่อพันธุ์ไทยของ	ชื่อพันธุ์ไทย	GSNumber	มีชื่อพันธุ์ไทย	-
2.	ชื่อพันธุ์อังกฤษของ	ชื่อพันธุ์อังกฤษ	GSNumber	มีชื่อพันธุ์ อังกฤษ	-
3.	มีกลิ่นหอม	GSNumber	เมล็ดข้าว	-	✓
4.	มีการนวด	GSNumber	เมล็ดมีเปลือก	-	✓
5.	การยี้ดของคอรวง	GSNumber	ต้น	-	✓
6.	การร่วงของเมล็ด	GSNumber	เมล็ดมีเปลือก	-	✓
7.	มีการเป็นท้องไข	GSNumber	เมล็ดข้าว	-	✓
8.	มีการแก่ของใบ	GSNumber	ใบ	-	✓

ในการออกแบบคุณสมบัติของตัวแบบออนโทโลยีข้าวไทยแบ่งออกเป็น 2 ลักษณะ ดังนี้ 1) คุณสมบัตិเชิงวัตถุ เป็นความสัมพันธ์ระหว่างคอนเซ็ปต์ ซึ่งคุณสมบัติดังกล่าวในตัวแบบออนโทโลยีข้าวไทยมีทั้งหมด 35 คุณสมบัติ ดังแสดงภาพที่ 4 และ 2) คุณสมบัติเชิงข้อมูล เป็นความสัมพันธ์ระหว่างคอนเซ็ปต์

และข้อมูล โดยคุณสมบัติดังกล่าวในตัวแบบออนโทโลยีข้าวไทยมีทั้งหมด 21 คุณสมบัติ ดังแสดงภาพที่ 5 ลักษณะของข้อมูลจะมีความแตกต่างกันขึ้นอยู่กับความเหมาะสมของข้อมูล อาทิ ข้อมูลชนิดข้อความ เลขจำนวนเต็ม เลขทศนิยม เป็นต้น



ภาพที่ 4 โครงสร้างคุณสมบัติเชิงวัตถุของตัวแบบออนโทโลยีข้าวไทย



ภาพที่ 5 โครงสร้างคุณสมบัติเชิงข้อมูลของตัวแบบออนโทโลยีข้าวไทย

วิจารณ์และสรุปผล

การประเมินการวัดประสิทธิภาพการค้นคืนในการสืบค้นข้อมูลจากข้อมูลเชิงความหมาย โดยการทำงานเริ่มทำงานเมื่อมีการส่งข้อมูลหรือคำสำคัญ (keyword) เข้าสู่ออนโทโลยี จากนั้นก็จะนำส่งข้อมูลหรือคำสำคัญไปค้นหาไปเปรียบเทียบและกับความรู้ในฐานองค์ความรู้ที่ได้ทำการเชื่อมโยงข้อมูลไว้แล้ว ซึ่งในที่นี้ประเมินการทำงานออนโทโลยีข้าวไทยได้ใช้ค่าความแม่นยำ วัดจากผลลัพธ์จัดกลุ่มข้อมูล (classification) ค่าของผลลัพธ์ที่ได้จากการจัดกลุ่ม [21, 22] สามารถแสดงดังตารางที่ 3

จากค่าผลลัพธ์ที่ได้ทั้ง 4 ค่า ในตารางที่ 3 สามารถนิยามการวัดประสิทธิภาพด้วยค่าความแม่นยำ (precision) ซึ่งเป็นค่าของการทำนายค่าที่จริงในกลุ่มที่

พบและถูกต้อง ซึ่งพิจารณาจากจำนวนกลุ่มข้อมูลทั้งหมด เมื่อ $(tp + fp)$ มีค่าเท่ากับ 0 ให้ตัวหาร $(tp + fp)$ มีค่าเป็น 1 [15, 16]

จากการทดลองนี้เป็นการทดลองการออกแบบออนโทโลยีข้าวไทยเพื่อช่วยในการสืบค้นคำสำคัญที่ได้ความสัมพันธ์ในการค้นคืนเชิงโครงสร้างและข้อมูลด้วยเทคโนโลยีออนโทโลยี ผลของการทดลองทั้งหมด 30 ครั้งที่ได้ตรงกับวัตถุประสงค์ที่กำหนดไว้จึงสามารถสืบค้นได้ตรงตามที่ต้องการด้วยค่าเฉลี่ยของความแม่นยำได้ถึง 76.29% (ตารางที่ 4) ดังนั้นสิ่งที่ควรพัฒนาถัดไปคือ นำสิ่งได้ออกแบบไว้มาพัฒนาเพื่อให้สามารถนำไปใช้ได้จริงสำหรับการค้นหาเชิงความหมายในอนาคต

ตารางที่ 3 ค่าของคอนฟิวชันเมทริกซ์ (confusion matrix) แบบ 2 กลุ่ม

ค่าที่แท้จริง (actual class)	ค่าที่ทำนายได้ (predicted class)		สมการค่าความแม่นยำ (precision)
	คลาส YES	คลาส NO	
คลาส YES	True Positive : tp (ผลลัพธ์ต้องมีแล้วถูกดึงออกมา)	False Negative : fn (ผลลัพธ์ต้องมีแต่ไม่ถูกดึงออกมา)	$\text{precision} = \frac{tp}{tp + fp}$
คลาส NO	False Positive: fp (ผลลัพธ์ต้องไม่มีแต่ถูกดึงออกมา)	True Negative : tn (ผลลัพธ์ต้องไม่มีและไม่ถูกดึงออกมา)	

ตารางที่ 4 ผลลัพธ์จากการค้นคืนเชิงโครงสร้างและข้อมูล

จำนวนครั้งที่ค้นคืน ข้อมูล	ผลลัพธ์ที่ต้องมีแล้ว ถูกดึงออกมา	ผลลัพธ์ต้องไม่มี แต่ถูกดึงออกมา	ผลลัพธ์ต้องมีแต่ ไม่ถูกดึงออกมา	ค่าความแม่นยำ
1	1	0	0	1.00
2	1	0	0	1.00
3	1	0	0	1.00
4	37	0	0	1.00
5	430	4	0	0.99078
...	...	...	...	...
30	0	0	0	0
ผลรวม				22.88639
$\text{ค่าเฉลี่ยความแม่นยำ} = \frac{22.88639}{30} \times 100 = 76.28797$				

กิตติกรรมประกาศ

งานวิจัยชิ้นนี้เป็นส่วนหนึ่งของโครงการศึกษาการสร้างต้นแบบออนโทโลยีขององค์ความรู้ข้าวไทย (Thai Rice Knowledge Ontology Development Prototyping) ซึ่งได้รับทุนอุดหนุนงานวิจัยจากสำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.)

เอกสารอ้างอิง

1. ธนิตา วงศ์กาฬสินธุ์, งามนิจ อาจอินทร์. การเชื่อมโยงออนโทโลยีบนโดเมน E-learning โดยใช้ WordNet บนพื้นฐานของการวัดความคล้ายคลึงในเชิงความหมาย. การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 5 (NCCIT2009); 2552. p.218-224.
2. เพ็ญพรรณ อัสวณพเกียรติ, กฤษณะ ไวยมัย. ระบบสำหรับสืบค้นและรวบรวมข้อมูลทางชีววิทยาโดยใช้ชีวออนโทโลยี. [วิทยานิพนธ์. วศ.ม. (วิศวกรรมคอมพิวเตอร์)], มหาวิทยาลัยเกษตรศาสตร์; 2547.
3. มูลนิธิข้าวไทย. จุดกำเนิดและประวัติข้าวไทย; 2006. [Internet]. [updated 2012 May 17], ได้

จาก http://www.thairice.org/html/aboutrice/about_rice1.htm.

4. ศิวกร สิริวัฒนภักษ์. การสร้างต้นแบบออนโทโลยีในประเทศไทยด้วยวิกิพีเดีย. [สารนิพนธ์ วิทยาการคอมพิวเตอร์], มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ; 2551
5. สถาบันวิจัยข้าว. ข้าว. กรมวิชาการเกษตร. กระทรวงเกษตรและสหกรณ์; 2547
6. สถาบันวิจัยข้าว. ฐานข้อมูลเชื้อพันธุ์พืช : ข้าว. กรมวิชาการเกษตร. กระทรวงเกษตรและสหกรณ์; 2544
7. สิตลา วงศ์กาฬสินธุ์ และงามนิจ อาจอินทร์. การรวมกันของออนโทโลยีอย่างมีความหมายสำหรับโดเมนห้องสมุดอิเล็กทรอนิกส์. การประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 5 (NCCIT2009); 2552. p.288-293.
8. สิริรัตน์ ประกฤติทาการชัย. การสร้างต้นแบบออนโทโลยีของพืชสมุนไพรไทย. [สารนิพนธ์, วิทยาการคอมพิวเตอร์], มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ; 2550
9. Antoniou G, Van Harmelen F. A Semantic Web Primer. Cambridge, MA, USA: The MIT Press; 2004

10. Berners-Lee T, Hendler J, and Lassila O. The Semantic Web. Scientific American; 2001. p.34 – 43.
11. Daconta MC, Obrst LJ, Smith KT. The Semantic Web—A Guide to the Future of XML, Web Services, and Knowledge Management. New York: Wiley; 2003.
12. Dodds L. Introducing SPARQL: Querying the Semantic Web; 2005. [cited 2012 May 10], Available from: <http://www.xml.com/pub/a/2005/11/16/introducing-sparql-querying-semantic-web-tutorial.html>.
13. Gruber TR. A Translation Approach to Portable Ontology Specifications. In: Stanford University. Knowledge Acquisition, Volume 5, Issue 2; 1993. p.199-220.
14. Lacy LW. OWL: Representing information using the web ontology language. Crewe: Trafford Publishing; 2005
15. Makhoul J, Kubala F, Schwartz R, Weischedel R. Performance measures for information extraction, In: Morgan Kaufmann. Proceedings of the DARPA Broadcast News Workshop; 1999. p.249-252
16. Maynard D, Peters, W, Li Y. Metrics for evaluation of ontology based information extraction. In: Proceedings of the 4th Workshop on Evaluating Ontologies for the Web; 2006.
17. Protégé. PROTEGE OVERVIEW; 2013. [cited 2012 April 2], Available from: <http://protege.stanford.edu/>.
18. Prud'hommeaux, E, Seaborne A. SPARQL query language for RDF; (2008). [cited 2012 April 5], Available from: <http://www.w3.org/TR/rdf-sparql-query>.
19. Sirin E, Parsia B. SPARQL-DL: SPARQL Query for OWL-DL. In: Proceedings of the OWLED 2007 Workshop on OWL: Experiences and Directions, Innsbruck, Austria; 2007
20. Song MH, Lim SY, Kang DJ, Lee SJ. Automatic Classification of Web Pages based on the Concept of Domain Ontology. In: Proceedings of the 12th Asia-Pacific Software Engineering Conference (APSEC05); 2005. p.645-654.
21. W3C. Resource Description Framework (RDF) Schema Specification 1.0; 2000. [cited 2012 April 2], Available from: <http://www.w3.org/TR/2000/CR-rdf-schema-20000327>.
22. W3C. RDF Primer; 2004. [cited 2012 April 2], Available from: <http://www.w3.org/TR/2004/REC-rdf-primer-20040210>.