



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (วิทยาการคอมพิวเตอร์)

ปริญญา

วิทยาการคอมพิวเตอร์

วิทยาการคอมพิวเตอร์

สาขา

ภาควิชา

เรื่อง การใช้ออนโทโลยีเพื่อการค้นคืนคำพิพากษาศาลฎีกา

Supreme Court Sentence Retrieval using Ontology

นามผู้วิจัย นายธนพล ตันติศรีปรีชา

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์นวลวรรณ สุนทรภิชช์, Ph.D.)

หัวหน้าภาควิชา

(ผู้ช่วยศาสตราจารย์ศิริกร จันทร์นวล, M.Sc.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญญา วีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ เดือน พ.ศ.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

การใช้ออนโทโลยีเพื่อการค้นคืนคำพิพากษาศาลฎีกา

Supreme Court Sentence Retrieval using Ontology

โดย

นายธนพล ตันติศรีปรีชา

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)

พ.ศ. 2553

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ธนพล ตันติศรีปรีชา 2553: การใช้ออนโทโลยีเพื่อการค้นคืนคำพิพาทภาษาศาสตร์
ปริญญาวิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์) สาขาวิทยาการคอมพิวเตอร์
ภาควิชาวิทยาการคอมพิวเตอร์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: ผู้ช่วยศาสตราจารย์
นวลวรรณ สุนทรภิชช์, Ph.D. 119 หน้า

งานวิจัยนี้มีจุดประสงค์เพื่อเสนอขั้นตอนวิธีการค้นคืนคำพิพาทภาษาศาสตร์ให้มีความ
หลากหลายในประเด็นทางกฎหมาย (SCRO II) ขั้นตอนวิธี SCRO ใช้หลักความสัมพันธ์ของ
Ontology ในการขยายคำค้น เพื่อให้ได้กลุ่มคำค้นที่นำไปค้นคืนเอกสาร จากนั้นเอกสารที่ได้จะถูก
จัดเรียงน้ำหนักตามค่า TFIDF ข้อมูลที่ใช้ในการทดลองของงานวิจัยนี้คือคำพิพาทภาษาศาสตร์ที่
เกี่ยวข้องกับคดีมรดกและตัวเงิน ในช่วงปี พ.ศ. 2542 ถึง พ.ศ. 2552 วิธีการขยายคำค้นที่ศึกษามี 3
แบบ คือ SCRO, SCRO I และ SCRO II ซึ่งมีความแตกต่างของโครงสร้างของออนโทโลยี และ
กลุ่มของคำค้น ประสิทธิภาพของขั้นตอน ประเมินโดยเปรียบเทียบค่าความหลากหลายของ
ประเด็นทางกฎหมาย ระหว่างขั้นตอนวิธี SCRO และขั้นตอนค้นคืนคำพิพาทภาษาศาสตร์ใน
ปัจจุบัน ผลการศึกษาพบว่าขั้นตอนวิธี SCRO II ซึ่งมีกลุ่มของคำค้นเป็นกลุ่มย่อยและใช้ออน
โทโลยี 2 ชั้น มีประสิทธิภาพในการค้นคืนคำพิพาทภาษาศาสตร์มากกว่าวิธีปัจจุบัน และให้ค่าความ
หลากหลายของประเด็นทางกฎหมายสูงสุดบนชุดข้อมูลมรดกและตัวเงินคิดเป็น 72% และ 79%
ตามลำดับ

ลายมือชื่อนิสิต

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Tanapon Tantisripreecha 2010: Supreme Court Sentence Retrieval using Ontology.
Master of Science (Computer Science), Major Field: Computer Science, Department of
Computer Science. Thesis Advisor: Assistant Professor Nuanwan Soonthornphisaj,
Ph.D. 119 pages.

This study presents an algorithm for Supreme Court Sentences Retrieval using Ontology (SCRO). SCRO applies relations of ontology to do query expansion, then; the set of queries are used to retrieve required documents. TFIDF is used to rank the weight of retrieved documents in respect of accuracy. Supreme Court Sentences of Succession Law and Bill of exchange from 1999 - 2009 are used in this study, SCRO, SCRO I and SCRO II are query expansion methods using different in ontology structure and set of queries. Performance of the methods are evaluated by comparing diversified law issues between SCRO methods and traditional methods. The experimental results show that using 2-level ontology with a subset of queries, SCRO II, can enhance the retrieval performance based on diversity value. The diversity coverage measure on Succession law and Bill of exchange law data sets are 72% and 79% respectively.

Student's signature

Thesis Advisor's signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ดี ด้วยความช่วยเหลืออย่างดียิ่งของ ผศ.ดร. นวลวรรณ
สุนทรภิชช์ ประธานกรรมการที่ปรึกษา ที่ได้สั่งสอนให้คำแนะนำ วิธีการและขั้นตอนในการ
ดำเนินงาน ซึ่งแนะแนวทางแก้ไขปัญหาและตรวจแก้ข้อบกพร่องต่างๆ ในงานวิจัยมาโดยตลอด

ขอกราบขอบพระคุณ คุณพ่อประพัทธ์ ดันติศรีปรีชา คุณแม่อรณา ดันติศรีปรีชา ที่ได้ให้
กำลังใจ ให้การสนับสนุน และความช่วยเหลืออย่างที่สุดมาโดยตลอดจนสำเร็จการศึกษา

ขอขอบพระคุณคณาจารย์ทุกท่านที่ให้การอบรม สั่งสอนวิชาความรู้ต่างๆมาจนสำเร็จ
การศึกษา

ขอขอบคุณบัณฑิตวิทยาลัย และคณะวิทยาศาสตร์ ที่ได้ให้ทุนสนับสนุนการประชุม
วิชาการที่สาธารณรัฐประชาธิปไตยประชาชนจีน

ขอขอบคุณภาควิชาวิทยาการคอมพิวเตอร์ ที่ได้สนับสนุนด้านอุปกรณ์ และสถานที่ตลอด
ระยะเวลาในการทำวิทยานิพนธ์

ขอขอบคุณพี่พิชญา ดันติศรีปรีชา พี่ชรินตรา ดันติศรีปรีชา พี่วิสิทธิ์ บุญชุม และพี่อณิมา
รอดเสียงล้ำ ที่ช่วยตรวจเอกสาร แนะนำการเขียนวิทยานิพนธ์ รวมทั้งช่วยฝึกซ้อมก่อนการสอบปาก
เปล่าครั้งสุดท้าย

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์นี้ ขอมอบแด่ผู้มีพระคุณทุกท่าน

ธนพล ดันติศรีปรีชา

เมษายน 2553

สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(4)
คำนำ	1
วัตถุประสงค์	2
การตรวจเอกสาร	3
อุปกรณ์และวิธีการ	20
อุปกรณ์	20
วิธีการ	20
ผลและวิจารณ์	46
ผล	46
วิจารณ์	74
สรุปและข้อเสนอแนะ	76
สรุป	76
ข้อเสนอแนะ	77
เอกสารและสิ่งอ้างอิง	78
ภาคผนวก	81
ภาคผนวก ก รายละเอียดคอนโทโลยี	82
ภาคผนวก ข ผลงานตีพิมพ์	97
ประวัติการศึกษา และการทำงาน	119

สารบัญตาราง

ตารางที่		หน้า
1	สรุปงานตามลำดับเวลางานวิจัย	19
2	ค่าสัมประสิทธิ์ของความสัมพันธ์	30
3	ขั้นตอนวิธี SCRO I	35
4	ขั้นตอนวิธี SCRO II	44
5	ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของประเด็นทาง กฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลมรดก	48
6	ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของ ขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลมรดก	49
7	ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของขั้นตอนวิธี Deka I และ Deka II บนชุดข้อมูลมรดก	50
8	ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของ ขั้นตอนวิธี Deka II และ Deka I บนชุดข้อมูลมรดก	51
9	การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความ หลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลมรดก	52
10	การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความ หลากหลายของประเด็นทางกฎหมาย ของขั้นตอนวิธี Deak I และ Deka II บนชุดข้อมูลมรดก	53
11	ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของขั้นตอนวิธี SCRO II, SCRO I และ Baseline บนชุดข้อมูลตัวเงิน	54
12	ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของ ขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลตัวเงิน	55
13	ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของขั้นตอนวิธี Deka I และ Deka II บนชุดข้อมูลตัวเงิน	56
14	ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของ ขั้นตอนวิธี Deka I และ Deka II บนชุดข้อมูลตัวเงิน	57

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
15	การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลตัวเงิน	58
16	การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความหลากหลายของประเด็นทางกฎหมาย ของขั้นตอนวิธี Dek a I และ Dek a II บนชุดข้อมูลตัวเงิน	59
17	ผลการทดลองที่ 2 บนชุดข้อมูลมรดก	64
18	ผลการทดลองที่ 2 บนชุดข้อมูลตัวเงิน	65
19	การทดลองที่ 2 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ และค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline และ SCRO I บนชุดข้อมูลมรดก	66
20	การทดลองที่ 2 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ และค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO I และ Baseline บนชุดข้อมูลตัวเงิน	67
21	คำค้นจากผู้ใช้และกลุ่มคำค้นที่เพิ่มขึ้นจากขั้นตอนวิธี SCRO	69
22	การทดลองที่ 3 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ และค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO และ Baseline บนชุดข้อมูลตัวเงิน	71
23	ผลการประเมินจากผู้เชี่ยวชาญ	72

ตารางผนวกที่

ก1	คำที่ใช้ในการสร้างออนโทโลยีจำนวน 260 คำ	83
ก2	คำในโนดของออนโทโลยีมรดกที่ใช้ในขั้นตอนวิธี SCRO II	93
ก3	คำในความสัมพันธ์ของออนโทโลยีมรดกที่ใช้ในขั้นตอนวิธี SCRO II	94
ก4	คำในโนดของออนโทโลยีตัวเงินที่ใช้ในขั้นตอนวิธี SCRO II	95
ก5	คำในความสัมพันธ์ของออนโทโลยีตัวเงินที่ใช้ในขั้นตอนวิธี SCRO II	96

สารบัญภาพ

ภาพที่		หน้า
1	ออนโทโลยีที่โน้ตมีความหมายเหมือนกัน	4
2	ออนโทโลยีที่โน้ตมีความหมายสัมพันธ์กัน	4
3	วิธีการเลือกคำแบบ Global Analysis	7
4	วิธีการเลือกคำแบบ Local Analysis	8
5	วิธีการเลือกคำแบบ Local Context Analysis	9
6	การใช้เวกเตอร์สเปซแทนเอกสาร	10
7	ขั้นตอนการขยายคำค้นโดยใช้ออนโทโลยี	12
8	กราฟ RDF triple ที่มีความสัมพันธ์โดยตรง	16
9	ระบบสืบค้นคำพินิจาและคำสั่งคำร้องศาลฎีกา	21
10	โครงสร้างของออนโทโลยีสำหรับขั้นตอนวิธี SCRO	24
11	ออนโทโลยีมรดกของขั้นตอนวิธี SCRO	25
12	ออนโทโลยีตัวเงินของขั้นตอนวิธี SCRO	26
13	ออนโทโลยีตัวเงินสำหรับการทดลองที่ 3	27
14	ขั้นตอนการทำงานของขั้นตอนวิธี SCRO	28
15	โครงสร้างออนโทโลยีสำหรับขั้นตอนวิธี SCRO I	31
16	ออนโทโลยีมรดกสำหรับขั้นตอนวิธี SCRO I	32
17	ออนโทโลยีตัวเงินสำหรับขั้นตอนวิธี SCRO I	33
18	แผนภาพขั้นตอนวิธี SCRO I และขั้นตอนวิธี SCRO II	34
19	โครงสร้างออนโทโลยี SCRO I	36
20	โครงสร้างออนโทโลยีสำหรับขั้นตอนวิธี SCRO II	38
21	ออนโทโลยีมรดกครั้งที่ 1	40
22	ออนโทโลยีมรดกครั้งที่ 2	41
23	ออนโทโลยีตัวเงินครั้งที่ 1	42
24	ออนโทโลยีตัวเงินครั้งที่ 2	42
25	ค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II	60

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
26	ค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Deka I และ Deka II	61
27	การค้นคืนคำพิพากษาที่แตกต่างกันในแต่ละกลุ่มคำค้นของขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline	70
28	เปรียบเทียบประสิทธิภาพระหว่างขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline	72

การใช้ออนโทโลยีเพื่อการค้นคืนคำพิพากษาศาลฎีกา

Supreme Court Sentence Retrieval using Ontology

คำนำ

ออนโทโลยี (Ontology) คือ โครงสร้างข้อมูลชนิดกราฟที่ใช้แทนความรู้เฉพาะด้าน (domain specific knowledge) ในรูปแบบโครงข่ายความหมาย ซึ่งใช้อธิบายความสัมพันธ์ระหว่างคำแต่ละคำว่ามีความสัมพันธ์กับคำอื่นอย่างไร โครงสร้างของออนโทโลยีประกอบด้วยโนดและความสัมพันธ์ โดยโนดคือคำที่ใช้เป็นตัวแทนหรือแนวความคิดของความรู้ และความสัมพันธ์คือเส้นเชื่อมระหว่างโนดตั้งแต่ 2 โนดขึ้นไปที่มีความสัมพันธ์กัน ความสัมพันธ์ดังกล่าวอาจเป็นคำที่มีความหมายเหมือนกัน คำที่มีความหมายตรงข้ามกัน หรือคำที่มีความสัมพันธ์กัน

ออนโทโลยีนี้นิยมนำมาใช้ในการค้นคืนเชิงความหมาย เพราะไม่เพียงแต่ค้นคืนคำที่มีความหมายเหมือนกับคำค้นหลักเพียงอย่างเดียว แต่ยังสามารถช่วยในการเลือกคำที่มีความสัมพันธ์กับคำค้นหลักในรูปแบบต่างๆ เช่น ความหมายเหมือนกัน ความหมายตรงข้ามกัน มีความสัมพันธ์กัน เป็นต้น เพิ่มเข้าไปยังกลุ่มของคำค้น ทำให้สามารถค้นคืนเอกสารได้ครอบคลุมและตรงกับความต้องการของผู้ใช้มากขึ้น

การศึกษาทางด้านกฎหมายต้องการคำพิพากษาศาลฎีกาที่มีความหลากหลายทางประเด็นของกฎหมาย เพื่อใช้ในการเรียนรู้สำหรับบุคคลที่ต้องการเปรียบเทียบข้อเท็จจริงต่างๆ ของกฎหมาย ซึ่งในปัจจุบันการสืบค้นคำพิพากษาศาลฎีกาใช้คำค้นหลักจากผู้ใช้เพียงอย่างเดียว หากผู้ใช้เป็นผู้ที่ไม่มีความรู้ในศัพท์เฉพาะทางด้านกฎหมายดีพอ อาจทำให้คำพิพากษาศาลฎีกาที่ค้นคืนได้ไม่มีความหลากหลายในแง่ของประเด็นทางกฎหมาย งานวิจัยนี้ได้เสนอขั้นตอนการสืบค้นคำพิพากษาศาลฎีกา โดยใช้ออนโทโลยีช่วยในการขยายคำสำคัญให้กับคำค้นหลักและกำหนดค่าความสัมพันธ์ระหว่างคำ เพื่อใช้ในการคำนวณค่าน้ำหนักของคำสำคัญที่ปรากฏในแต่ละคำพิพากษาศาลฎีกา ทำให้ได้ประเด็นทางกฎหมายจากคำพิพากษาศาลฎีกาที่มีความหลากหลาย

วัตถุประสงค์

เพื่อเสนอขั้นตอนวิธีในการค้นคืนคำพิพากษาศาลฎีกาให้มีประเด็นทางกฎหมายที่มีความหลากหลาย

ประโยชน์ที่คาดว่าจะได้รับ

สามารถค้นคืนคำพิพากษาศาลฎีกาเพื่อให้ได้กลุ่มของฎีกาที่มีประเด็นทางกฎหมายที่หลากหลาย

ขอบเขตและข้อจำกัด

ประเด็นทางกฎหมายที่พิจารณาในงานวิจัยนี้เกี่ยวข้องกับกฎหมายในส่วนสารบัญญัติตามประมวลกฎหมายแพ่งและพาณิชย์ว่าด้วยเรื่องตัวเงินมาตรา 898-1001 และ คำพิพากษาศาลฎีกาว่าด้วยเรื่องมรดกมาตรา 1599-1754 เท่านั้น

การตรวจเอกสาร

ทฤษฎีที่เกี่ยวข้องกับงานวิจัย

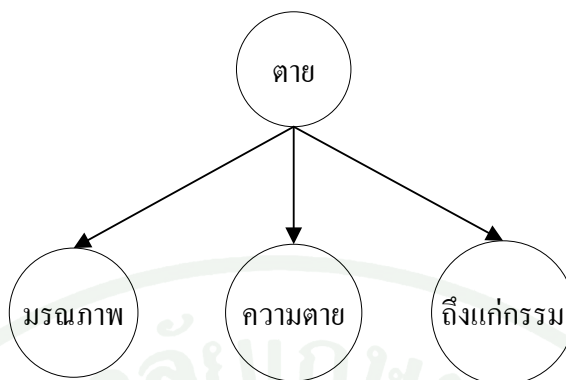
1. ออนโทโลยี

ออนโทโลยีคือโครงสร้างข้อมูลชนิดกราฟที่ใช้แทนความรู้เฉพาะด้านในรูปแบบโครงข่ายความหมาย (Gruber., 1993) ซึ่งใช้อธิบายความสัมพันธ์ระหว่างคำแต่ละคำที่มีความสัมพันธ์กับคำอื่นๆ อย่างไร โครงสร้างของออนโทโลยี ประกอบด้วย โหนดและความสัมพันธ์

โหนดคือคำที่ใช้เป็นตัวแทน หรือแนวความคิด (Concept) ของเอกสาร โดยคำที่ใช้ส่วนใหญ่เป็นคำนาม เนื่องจากเป็นคำที่มีความหมายในตัวเอง ไม่ต้องการคำที่มาเป็นส่วนขยายอีก ซึ่งโหนดในออนโทโลยีสามารถประกอบด้วยคำอื่นที่ไม่ใช่คำนามได้ เช่น คำกริยา วิเศษณ์ เป็นต้น

ความสัมพันธ์คือเส้นเชื่อมระหว่างโหนด 2 โหนดขึ้นไป ที่มีความสัมพันธ์กัน โดยในเส้นเชื่อมอาจมีคำที่ใช้บอกความสัมพันธ์ เช่น คำที่มีความหมายเหมือนกัน ความหมายตรงข้ามกัน หรือคำที่มีความสัมพันธ์กัน เป็นต้น ความสัมพันธ์ที่นิยมใช้ในออนโทโลยี เช่น “คือ” (Is-a), “ส่วนหนึ่งของ” (Part-of)

ภาพที่ 1 แสดงตัวอย่างของออนโทโลยีที่โหนดมีความหมายเหมือนกัน โดยที่คำว่า “ตาย” เป็นโหนดแม่ และมีโหนดลูกจำนวน 3 โหนด ที่มีความหมายเหมือนกับโหนดแม่คือคำว่า “มรณภาพ” “ความตาย” “ถึงแก่กรรม” โดยมีความหมายว่า “ตาย” เช่นกัน



ภาพที่ 1 ออนโทโลยีที่โนดมีความหมายเหมือนกัน

ภาพที่ 2 ออนโทโลยีที่โนดมีความหมายสัมพันธ์กัน เช่น คำว่า “ผู้ตาย” มีความหมายสัมพันธ์กับคำว่า “มรคก” และมีความสัมพันธ์กับคำว่า “ทายาท”



ภาพที่ 2 ออนโทโลยีที่โนดมีความหมายสัมพันธ์กัน

ขั้นตอนการกำหนดโนดและความสัมพันธ์ในออนโทโลยี มี 2 วิธี ดังนี้

1. สร้างจากผู้เชี่ยวชาญคือให้ผู้เชี่ยวชาญในแต่ละด้านเป็นผู้กำหนดคำต่างๆ ให้กับโนดและความสัมพันธ์ของออนโทโลยี ทำให้ได้ออนโทโลยีที่มีความถูกต้อง และสมบูรณ์ แต่วิธีนี้ใช้เวลานาน และผู้เชี่ยวชาญอาจกำหนดความสัมพันธ์ของคำได้ไม่ครบทุกคำ

2. สร้างโดยอัตโนมัติโดยนำเอกสารที่ใช้เป็นข้อมูลในการสร้างออนโทโลยีมาหาความสัมพันธ์ของคำโดยใช้พจนานุกรม หรือใช้หลักของภาษารธรรมชาติ (Natural Language Processing) ทำให้สามารถสร้างออนโทโลยีได้อย่างรวดเร็ว แต่มีข้อเสียคือต้องใช้ผู้เชี่ยวชาญในการพิจารณาความสัมพันธ์ในออนโทโลยี เพื่อให้ได้ออนโทโลยีที่ถูกต้อง

2. การขยายคำค้น (Query expansion)

การขยายคำค้นคือการเพิ่มกลุ่มคำค้นใหม่ให้กับคำค้นหลักที่รับมาจากผู้ใช้ ให้ได้กลุ่มของคำที่มีความสัมพันธ์กัน เนื่องจากการค้นคืนเอกสาร บางเอกสารที่ตรงกับความต้องการของผู้ใช้ อาจไม่มีคำค้นหลักปรากฏในเอกสาร แต่มีคำอื่นที่มีความหมายคล้ายกันปรากฏในเอกสาร วิธีการขยายคำค้นจะช่วยปรับปรุงผลลัพธ์ของการค้นคืนเอกสารได้ การขยายคำค้นอาจมาจากการขยายคำที่มีความสัมพันธ์กัน, คำที่มีความหมายเหมือนกัน, หรือคำที่มีความหมายตรงข้ามกัน โดยวิธีที่สามารถบอกความสัมพันธ์ระหว่างคำที่ช่วยในการขยายคำค้นได้ เช่น ออนโทโลยี พจนานุกรม เวิร์ดเน็ต (Soonhee., 2006) เป็นต้น

งานวิจัยของ (Wu *et al.*, 2007) รวบรวมวิธีการเลือกคำ ในขั้นตอนการขยายคำค้นสามารถแบ่งได้เป็น 3 กรณีดังต่อไปนี้

2.1 ผู้ใช้มีส่วนในการเลือกคำ

การขยายคำค้นโดยผู้มีส่วนในการเลือกคำ ระบบจะแสดงคำค้นใหม่ให้กับผู้ใช้ เพื่อให้ผู้ใช้เป็นผู้เลือกกลุ่มของคำค้นที่ใช้ในการค้นคืนข้อมูลต่อไป

ข้อดีคือทำให้สามารถค้นคืนข้อมูลได้ตรงกับความต้องการของผู้ใช้ เนื่องจากผู้ใช้เป็นผู้เลือกคำที่ต้องการเพิ่ม ทำให้คำที่ไม่ตรงกับความต้องการของผู้ใช้จะไม่นำไปเพิ่มในกลุ่มคำค้นใหม่

ข้อเสียคือไม่สามารถใช้กับการค้นคืนเรื่องเฉพาะได้ เนื่องจากคำที่เป็นเรื่องเฉพาะ เป็นคำที่ต้องอาศัยประสบการณ์ในการเรียนรู้ และใช้เวลาศึกษานาน ซึ่งคนทั่วไปจะไม่เข้าใจความหมายที่แท้จริงของสิ่งที่ต้องการค้นคืน ทำให้ไม่สามารถเพิ่มกลุ่มของคำค้นที่ถูกต้องได้

2.2 การใช้พจนานุกรมในการเลือกคำ

คือการใช้พจนานุกรมในการตัวอ้างอิงกับคำค้นหลัก โดยคำค้นที่ได้มาจากความหมายของคำค้นหลัก เช่น คำที่มีความหมายเหมือนกัน คำที่มีความหมายตรงกันข้าม เป็นต้น

ข้อดีคือสามารถเพิ่มคำค้นได้โดยอัตโนมัติ โดยใช้พจนานุกรมเป็นตัวอ้างอิงในคำที่ต้องการขยายคำค้น เหมาะสำหรับข้อมูลที่เป็นเรื่องทั่วไป

ข้อเสียคือไม่สามารถค้นคืนเอกสารที่เป็นเรื่องเฉพาะได้ เนื่องจากความหมายในพจนานุกรมเป็นเรื่องทั่วไป ซึ่งรวมคำหลายๆ คำไว้ โดยที่บางคำอาจไม่เกี่ยวข้อง ทำให้มีคำที่ไม่จำเป็นต่อการค้นคืน

2.3 การเลือกคำโดยใช้เอกสารเป็นตัวอ้างอิง

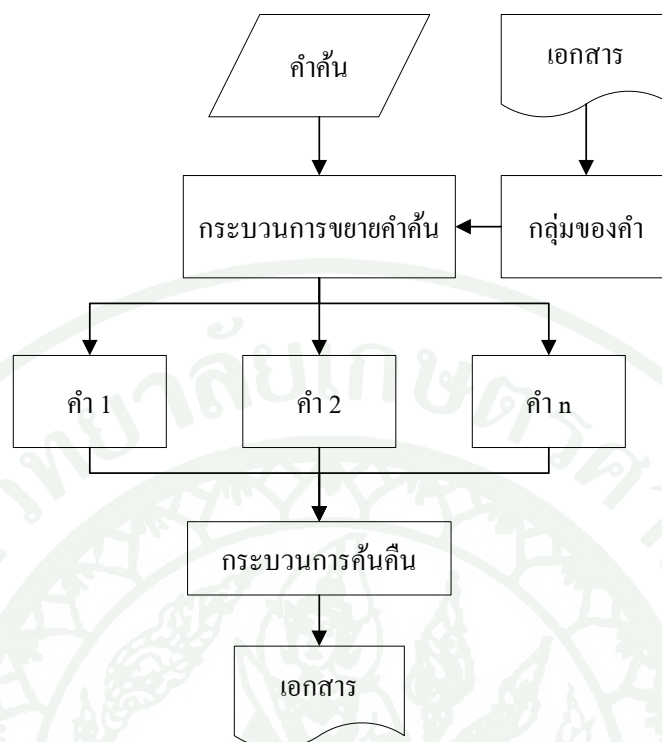
การเลือกคำโดยใช้เอกสารเป็นตัวอ้างอิงคือการขยายคำค้นโดยที่ใช้เอกสารที่ได้รวบรวมไว้ในด้านที่สนใจ มาใช้เลือกคำสำคัญเพื่อเพิ่มให้กับกลุ่มของคำค้นใหม่ โดยมีวิธีการเลือกคำค้นใหม่ ดังนี้

2.3.1 Global Analysis คือการสร้างกลุ่มของคำค้นที่มีความหมายเหมือนกัน จากเอกสารทั้งหมดในฐานข้อมูล โดยวิธีการเลือกคำจะพิจารณาว่าคำค้นหลักจากผู้ใช้เป็นคำค้นในกลุ่มใด ด้วยค่าการเกิดร่วมกัน (co-occurrence) และนำคำในกลุ่มนั้นมาใช้เป็นกลุ่มของคำค้น แสดงดังภาพที่ 3 ซึ่งมีรายละเอียด ดังต่อไปนี้

- 1) สร้างกลุ่มของคำค้น จากเอกสารทั้งหมดที่เก็บรวบรวมไว้ โดยคำที่มีความหมายเหมือนกันจะอยู่กลุ่มเดียวกัน และเก็บในรูปแบบเมทริกซ์
- 2) รับคำค้นจากผู้ใช้
- 3) พิจารณาคำค้นจากผู้ใช้ว่าควรอยู่ในกลุ่มใดจากค่าการเกิดร่วมกัน และดึงคำจากกลุ่มนั้นมาสร้างเป็นกลุ่มของคำค้นใหม่
- 4) นำกลุ่มของคำค้นใหม่ค้นคืนข้อมูล

ข้อดีคือสามารถดึงคำที่มีความหมายเหมือนกันได้ทุกคำ จากกลุ่มของคำค้นที่อยู่ในกลุ่มเดียวกัน

ข้อเสียคือต้องมีการจัดกลุ่มของคำจากเอกสารก่อนจึงสามารถค้นคืนได้ ทำให้ไม่สามารถทำงานได้แบบเวลาจริง (Real Time)



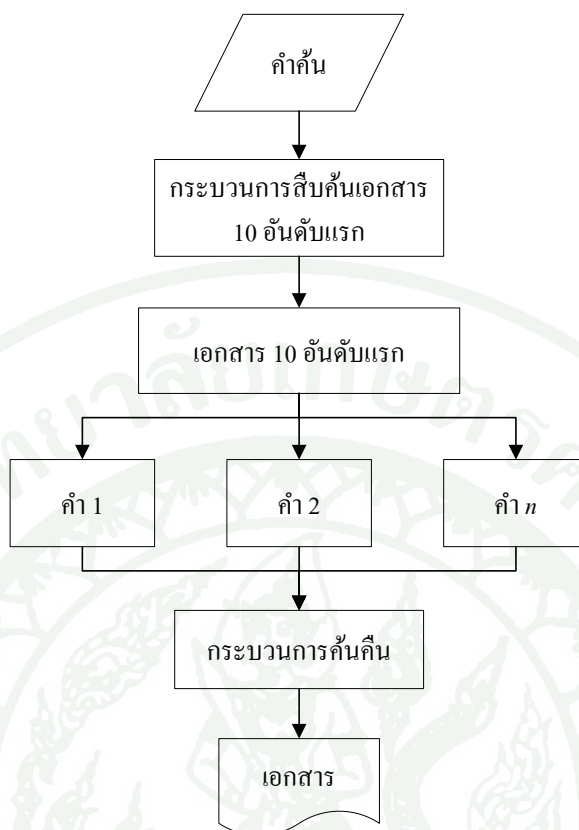
ภาพที่ 3 วิธีการเลือกคำแบบ Global Analysis

2.3.2 Local Analysis คือการขยายคำค้นโดยมีสมมุติฐานว่าเอกสาร 10 อันดับแรกที่ถูกค้นคืนคือเอกสารที่ตรงกับความต้องการของผู้ใช้ เนื่องจากเอกสารทั้ง 10 อันดับแรกมีคำที่ตรงกับความต้องการของผู้ใช้ ซึ่งขั้นตอนวิธีจะสกัดคำสำคัญจากเอกสารทั้ง 10 อันดับแรกออกมาเป็นกลุ่มของคำค้นใหม่ แสดงดังภาพที่ 4 ซึ่งมีรายละเอียดดังต่อไปนี้

- 1) รับคำค้นจากผู้ใช้
- 2) นำคำค้นจากผู้ใช้ ค้นคืนเอกสาร 10 อันดับแรก
- 3) เลือกคำสำคัญที่ปรากฏในเอกสาร 10 อันดับแรก เป็นกลุ่มของคำค้นใหม่
- 4) นำกลุ่มของคำค้นใหม่ค้นคืนข้อมูล

ข้อดีคือสามารถทำงานได้ในเวลาจริง เนื่องจากไม่ต้องมีการแบ่งกลุ่มของเอกสารที่ได้เตรียมไว้ แต่เป็นการดึงเอกสารออกมา 10 อันดับแรกจากคำค้นของผู้ใช้

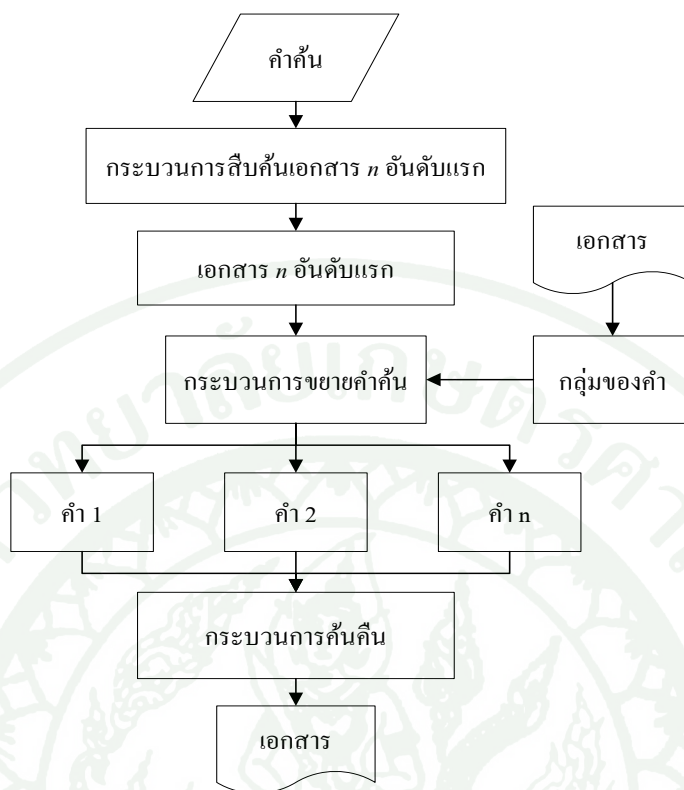
ข้อเสียคือถ้าเอกสารที่ค้นคืนได้ 10 อันดับแรก เป็นเอกสารที่ไม่ตรงกับความต้องการ ทำให้การค้นคืนจากคำที่เพิ่มขึ้นมาได้เอกสารที่ไม่ตรงกับความต้องการ



ภาพที่ 4 วิธีการเลือกคำแบบ Local Analysis

2.3.3 Local Context Analysis คือการขยายคำค้นที่ใช้หลักของ Global Analysis และ Local Analysis ร่วมกันคิดค้น โดย (Xu and Croft, 2000) โดยให้ n เอกสารแรกเป็นเอกสารที่ตรงกับความต้องการ และใช้ความสัมพันธ์ของการเกิดร่วมกันระหว่างกลุ่มของคำค้น แบ่งไว้กับเอกสาร n อันดับแรกในการเพิ่มคำค้นให้กับกลุ่มของคำค้นใหม่ แสดงดังภาพที่ 5 ซึ่งมีขั้นตอนดังต่อไปนี้

- 1) สร้างกลุ่มของคำค้น จากเอกสารทั้งหมดที่เก็บรวบรวมไว้ โดยคำที่มีความหมายเหมือนกันจะอยู่กลุ่มเดียวกัน
- 2) รับคำค้นจากผู้ใช้
- 3) นำคำค้นจากผู้ใช้ค้นคืนเอกสาร n อันดับแรก
- 4) เลือกคำสำคัญที่ปรากฏในเอกสาร n อันดับแรก และพิจารณาพร้อมกับกลุ่มของคำค้นที่ได้สร้างไว้
- 5) นำกลุ่มของคำค้นใหม่ไปค้นคืนข้อมูล



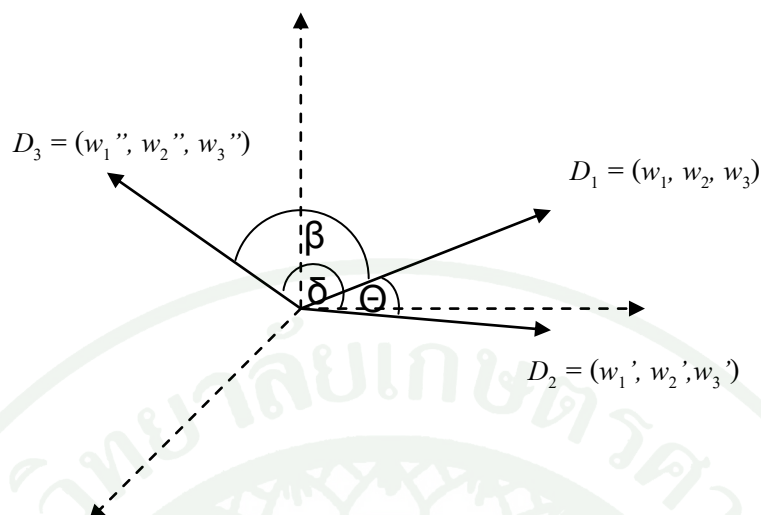
ภาพที่ 5 วิธีการเลือกคำแบบ Local Context Analysis

ข้อดีคือสามารถค้นคืนเอกสารได้ในเวลาจริง สามารถค้นคืนได้ทุกภาษา และไม่มีการเพิ่มคำค้นที่มีความกำกวม

ข้อเสียคือผู้ใช้ต้องระบุคำค้นที่เป็นคำกำกวมให้กับระบบ จึงสามารถทำงานได้อย่างถูกต้อง

3. เวกเตอร์สเปซโมเดล (Vector Space Model)

เวกเตอร์สเปซโมเดลคือวิธีคำนวณหาเอกสารที่เป็นตัวแทนกลุ่มของคำ โดยเอกสารแต่ละฉบับแทนด้วยเวกเตอร์ของคำ 1 เวกเตอร์ เพื่อใช้คำนวณหาความคล้ายคลึงกันของแต่ละเอกสาร โดยมีสมมติฐานว่าเวกเตอร์ของเอกสารที่มีความคล้ายกัน ย่อมทำมุมระหว่างกัน โดยมีขนาดของมุมน้อยกว่าเวกเตอร์ของเอกสารที่มีความแตกต่างกัน (Salton *et al.*, 1975) แสดงดังภาพที่ 6



ภาพที่ 6 การใช้เวกเตอร์สเปซแทนเอกสาร

จากภาพที่ 6 ค่าเวกเตอร์ของแต่ละเอกสารอยู่ในตำแหน่งที่ต่างกันในกราฟ 3 มิติ แสดงให้เห็นว่าเอกสาร D_1 มีความคล้ายคลึงกับเอกสาร D_2 มากกว่าเอกสาร D_3 เนื่องจาก มุมระหว่างเวกเตอร์ D_1 และ D_2 มีขนาดของมุมน้อยกว่าขนาดของมุมเวกเตอร์ D_1 และ D_3 ซึ่งการคำนวณค่าเวกเตอร์ของแต่ละเอกสาร สามารถคำนวณได้จากสมการที่ (1)

$$D_i = (w_{1i}, w_{2i}, \dots, w_{ni}) \quad (1)$$

โดย D_i คือ ค่าเวกเตอร์ของเอกสาร i
 w_{ni} คือ ค่าน้ำหนักของคำ
 i คือ ลำดับของเอกสาร
 n คือ ลำดับของคำ

ค่าเวกเตอร์ของเอกสาร i คือค่าของเวกเตอร์ที่ใช้แทนน้ำหนักเอกสาร โดยคำนวณจากค่าน้ำหนักของแต่ละคำ ซึ่งค่าน้ำหนักของแต่ละคำสามารถคำนวณจากสมการที่ (2)

$$w_{ij} = tf_{ij} * idf_j \quad (2)$$

โดย w_{ij} คือ ค่าน้ำหนักของคำ i ที่ปรากฏในคำพิพาท j
 tf_{ij} คือ ความถี่ของคำ i ที่ปรากฏในคำพิพาท j
 idf_j คือ ค่าที่ใช้วัดการกระจายของคำในแต่ละเอกสาร j

จากสมการที่ (2) ค่า w_{ij} คือค่าค่าน้ำหนักของคำแต่ละคำในเอกสาร โดยค่าค่าน้ำหนักของคำที่มีค่ามาก หมายถึงคำนั้นมีความสำคัญมากกว่าคำอื่นๆ คำนวณจากค่า $tf_{ij} * idf_j$ ซึ่ง tf_{ij} คือค่าความถี่ของคำที่เกิดขึ้นในเอกสารแต่ละเอกสาร สามารถคำนวณได้จากสมการที่ (3) และ idf_j คือค่าการกระจายของคำในแต่ละเอกสาร สามารถคำนวณได้จากสมการที่ (4)

$$tf_{i,j} = \frac{f_{i,j}}{\max(f_i)} \quad (3)$$

โดย $f_{i,j}$ คือ ความถี่ของคำ i ที่ปรากฏในคำพิพาทภาษา j
 $\max(f_i)$ คือ ค่าความถี่สูงสุดของคำ i

จากสมการที่ (3) ค่า tf_{ij} คือความถี่ของคำ โดยคำนวณได้จากความถี่ของคำในแต่ละเอกสาร และหารด้วยค่าสูงสุดของความถี่

$$idf_i = \ln\left(\frac{N}{1+n_i}\right) \quad (4)$$

โดย N คือ จำนวนเอกสารทั้งหมด
 n_i คือ จำนวนเอกสารที่มีคำ i ปรากฏอยู่

จากสมการที่ (4) ค่า idf_i คือค่าที่ใช้วัดการกระจายของคำในเอกสารต่างๆ โดยที่ N คือจำนวนเอกสารทั้งหมด n_i คือเอกสารที่มีคำ i ปรากฏอยู่ และ 1 คือค่าที่บวกเข้าไปเพื่อไม่ให้ตัวหารเป็นศูนย์

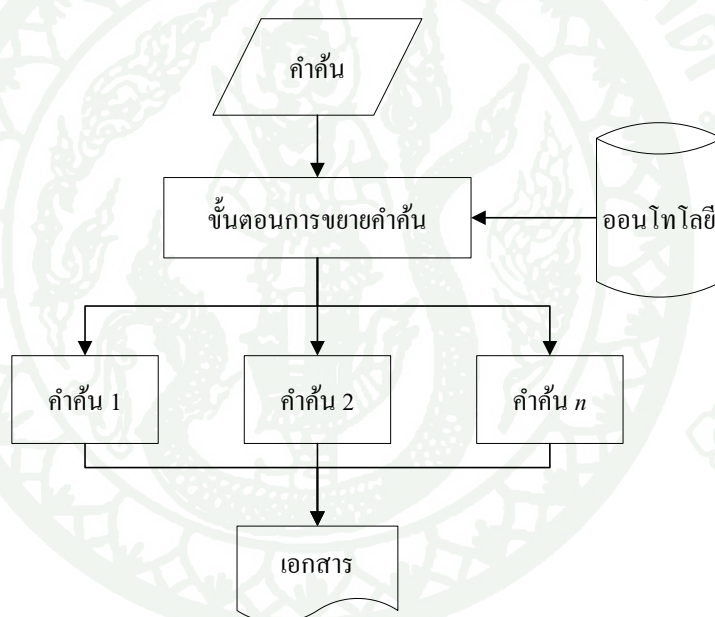
4. การขยายคำค้นโดยใช้ออนโทโลยี

การขยายคำค้นโดยใช้ออนโทโลยีคือการขยายคำค้นโดยอาศัยความสัมพันธ์ระหว่างคำที่เกิดขึ้นในออนโทโลยี การขยายคำค้นโดยใช้ออนโทโลยี จะดึงคำที่มีความหมายเหมือนกัน คำที่มีความหมายแตกต่างกัน รวมถึงคำที่มีความสัมพันธ์กับคำค้นหลัก แสดงดังภาพที่ 7 ซึ่งมีรายละเอียดดังต่อไปนี้

- 1) รับคำค้นจากผู้ใช้
- 2) สร้างกลุ่มของคำค้นโดยใช้ความสัมพันธ์ของออนโทโลยี
- 3) นำกลุ่มของคำค้นใหม่ค้นคืนเอกสาร

ข้อดีคือทำให้ได้กลุ่มของคำค้นใหม่ที่มีความสัมพันธ์กัน และสามารถขยายคำที่เป็นเรื่องเฉพาะได้

ข้อเสียคือต้องมีการสร้างออนโทโลยีของแต่ละเรื่อง โดยผู้เชี่ยวชาญในแต่ละด้าน ซึ่งในเวลานานในการสร้าง



ภาพที่ 7 ขั้นตอนการขยายคำค้นโดยใช้ออนโทโลยี

5. การวัดประสิทธิภาพ

การวัดประสิทธิภาพของการค้นคืนข้อมูลคือขั้นตอนในการตรวจสอบว่า เอกสารที่ค้นคืนได้นั้นตรงกับความต้องการของผู้ใช้หรือไม่ ถ้าเนื้อหาในเอกสารเป็นข้อมูลเฉพาะทาง เช่น คำพิพากษาศาลฎีกา ข้อมูลด้านการแพทย์ ข้อมูลด้านการเกษตร เป็นต้น จะต้องให้ผู้เชี่ยวชาญในแต่ละด้านเป็นผู้ตรวจสอบ แต่ถ้าเอกสารเป็นเรื่องทั่วไป สามารถให้ผู้ใช้เป็นผู้ตรวจสอบได้ ดังนั้นการวัดประสิทธิภาพของการค้นคืนเอกสารจึงมีวิธีการวัดประสิทธิภาพที่แตกต่างกันดังนี้

5.1 ผู้เชี่ยวชาญเป็นผู้วัดประสิทธิภาพของเอกสารที่ค้นคืนได้ โดยผู้เชี่ยวชาญพิจารณาว่าเอกสารแต่ละเอกสารที่ค้นคืนได้ เป็นเอกสารที่ตรงกับความต้องการหรือไม่ เนื่องจากข้อมูลบางประเภทเป็นเรื่องเฉพาะทางซึ่งต้องใช้ผู้เชี่ยวชาญเท่านั้นในการให้คำตอบกับผลของการค้นคืนเอกสาร

5.2 ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) เป็นค่าที่ใช้วัดประสิทธิภาพของการค้นคืนเอกสารที่นิยมมากที่สุด โดยแสดงดังสมการต่อไปนี้

5.2.1 ค่าความแม่นยำคือค่าที่ใช้วัดประสิทธิภาพของขั้นตอนวิธีการค้นคืน สามารถค้นคืนเอกสารที่ตรงกับความต้องการของผู้ใช้มากหรือน้อย จากเอกสารทั้งหมดที่ถูกค้นคืน จำนวนจากสมการที่ (5)

$$\text{ค่าความแม่นยำ} = \frac{\text{จำนวนเอกสารที่ถูกค้นคืนและตรงกับความต้องการของผู้ใช้}}{\text{จำนวนเอกสารทั้งหมดที่ถูกค้นคืน}} \quad (5)$$

5.2.2 ค่าความระลึกคือค่าที่ใช้วัดประสิทธิภาพวิธีการค้นคืน สามารถค้นคืนเอกสารที่ตรงกับความต้องการได้มากหรือน้อย จากเอกสารที่ตรงกับความต้องการทั้งหมด จำนวนได้จากสมการที่ (6)

$$\text{ค่าความระลึก} = \frac{\text{จำนวนเอกสารที่ถูกค้นคืนและตรงกับความต้องการของผู้ใช้}}{\text{จำนวนเอกสารทั้งหมดในชุดข้อมูลที่ตรงกับความต้องการของผู้ใช้}} \quad (6)$$

5.3 ค่าความหลากหลายของประเด็นทางกฎหมายคือค่าที่ใช้วัดความหลากหลายของผลลัพธ์ของคำพิพากษาศาลฎีกา จำนวนระหว่างจำนวนประเด็นความหลากหลายที่ค้นคืนได้ ส่วนจำนวนประเด็นความหลากหลายทั้งหมด ซึ่งผลลัพธ์จะแสดงความหลากหลายของแต่ละคำค้น ดังสมการที่ (7)

$$\text{ค่าความหลากหลายของประเด็นทางกฎหมาย} = \frac{\text{จำนวนประเด็นความหลากหลายที่ค้นคืนได้}}{\text{จำนวนประเด็นความหลากหลายทั้งหมด}} \quad (7)$$

งานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ได้เสนอขั้นตอนค้นคืนคำพินิจภาษาศัลยวิทยาโดยใช้ออนโทโลยีในการขยายคำค้นจากผู้ใช้งาน ผู้วิจัยได้แบ่งการศึกษางานวิจัยที่เกี่ยวข้องออกเป็น 2 ส่วนคืองานวิจัยเกี่ยวกับการค้นคืนและขยายคำค้น และงานวิจัยเกี่ยวกับการค้นคืนเอกสารด้านกฎหมาย

1. การค้นคืนและขยายคำค้น

งานวิจัยเกี่ยวกับการขยายคำค้น มีสาเหตุมาจากคำค้นหลักจากผู้ใช้งานเพียง 1 คำไม่สามารถค้นคืนเอกสารที่ตรงกับความต้องการของผู้ใช้ได้ทั้งหมด เนื่องจากเอกสารที่ต้องการไม่มีคำค้นปรากฏอยู่ แต่มีคำที่มีความหมายเหมือนกันแทน หรือเอกสารที่ต้องการเป็นเอกสารเฉพาะด้านซึ่งผู้ใช้ไม่สามารถป้อนคำค้นเข้าไปเพื่อค้นหาเอกสารที่ต้องการได้ การขยายคำค้นได้ถูกรวบรวมโดย (Xu *et al.*, 2000) ซึ่งได้เสนอขั้นตอนวิธี Local Context Analysis โดยนำข้อดีของขั้นตอนวิธี Global analysis คือการจัดกลุ่มของคำสำคัญก่อนขยายคำค้นและ Local analysis คือการสมมติว่าเอกสารที่ค้นคืนมา 10 อันดับแรกเป็นเอกสารที่ตรงกับความต้องการ มาประยุกต์เป็นวิธี Local Context Analysis คือการจัดกลุ่มของคำค้นจากเอกสารที่ค้นคืน 10 อันดับแรก และเลือกกลุ่มคำที่ตรงกับคำค้น โดยผลลัพธ์จากการขยายคำค้นจะช่วยลดคำที่ไม่จำเป็นออกจากคำค้นที่เพิ่มขึ้น

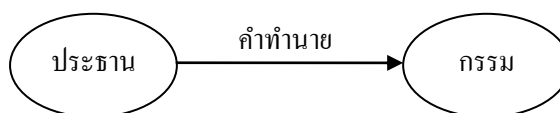
ปัญหาของการเลือกคำค้นใหม่คือคำบางคำอาจถูกขยายเพิ่มเข้ามาในกลุ่มของคำค้นมากเกินไป ส่งผลให้ในการค้นคืนเอกสาร ได้เอกสารที่ไม่ตรงกับความต้องการออกมาด้วย ทำให้ค่าความระลึก และค่าความแม่นยำมีค่าลดลง (Fiaidhi *et al.*, 2003) ได้พัฒนารูปแบบของเว็บไซต์สำหรับการค้นหาข้อมูล โดยใช้ออนโทโลยีเพื่อให้สามารถหาเอกสารที่ตรงกับความต้องการได้มากยิ่งขึ้น โดยใช้แนวคิดจากออนโทโลยีในการขยายคำค้นแทนรูปแบบการใช้คำค้นเข้าไปเปรียบเทียบกับฐานข้อมูล

ปัญหาของการค้นคืนเอกสารที่ได้ผลลัพธ์ไม่ตรงกับความต้องการของผู้ใช้ เกิดจากคำที่ใช้ในการค้นคืนไม่ปรากฏในเอกสารที่ผู้ใช้งานต้องการ คำค้นเหล่านั้นไม่สามารถสื่อความหมายหรือเชื่อมโยงความสัมพันธ์กับคำอื่นๆ และผู้ใช้ไม่สามารถระบุคำที่สามารถสื่อถึงเอกสารที่ตรงกับความต้องการได้ถูกต้อง ทำให้เกิดช่องว่างระหว่างคอมพิวเตอร์และมนุษย์ เนื่องจากคำที่ปรากฏในเอกสารนั้นเป็นคำที่มีความหมายเหมือนกับคำค้น และคำที่ใช้ค้นคืนเป็นคำที่สั้นเกินไป ซึ่ง

รายละเอียดไม่เพียงพอสำหรับการค้นคืนเอกสาร งานวิจัยของ (Xu *et al.*, 2007) ได้เสนอขั้นตอนการค้นคืนเอกสาร โดยใช้ออนโทโลยีในการขยายคำค้น ร่วมกับความถี่ของคำที่ปรากฏ โดยได้แบ่งความสัมพันธ์ของออนโทโลยีออกเป็น 2 ส่วนคือความสัมพันธ์ที่อยู่ในกลุ่มเดียวกัน และความสัมพันธ์ที่ใช้เป็นตัวอย่าง และมีฟังก์ชันในการกรองความสัมพันธ์โดยใช้วิธีการเปรียบเทียบการเกิดร่วมกันของคำที่ได้มา ก่อนนำคำไปค้นคืนเอกสาร ข้อมูลที่ใช้ในการทดลองคือบทความจาก yahoo.com จำนวน 82 บทความ โดยสามารถปรับปรุงประสิทธิภาพได้ 2 อย่างคือการขยายคำค้นจากออนโทโลยีใกล้เคียงกับเอกสารของผู้เชี่ยวชาญมากกว่าการขยายโดยใช้พจนานุกรม และคำจากการขยายเป็นคำพื้นฐานของกลุ่มเอกสารที่รวบรวมไว้เพิ่มขึ้น

งานวิจัยของ (Dridi and Ahmed., 2008) ได้นำเสนอขั้นตอนการค้นคืนเชิงความหมาย โดยสร้างออนโทโลยีที่แสดงถึงความสัมพันธ์ต่างๆ ของคำ โดยขั้นตอนของงานวิจัยนี้ได้ใช้การประมวลผลภาษาธรรมชาติในการวิเคราะห์ความหมาย และชนิดของคำ เพื่อนำมาสร้างเป็นออนโทโลยี โดยความสัมพันธ์ที่ใช้คือ “คือ” (is-a) และ “ส่วนหนึ่งของ” (part-of) คำที่ขยายออกมาจะมีค่าน้ำหนักไม่เท่ากัน เนื่องจากแต่ละคำมีความสำคัญไม่เท่ากัน จึงมีการปรับค่าน้ำหนักโดยใช้ค่า *TFIDF* ข้อมูลที่ใช้ในการทดลองคือเอกสารทางการแพทย์ด้านมะเร็งเต้านม ผลการทดลองทำการเปรียบเทียบการค้นคืนโดยใช้ออนโทโลยี กับขั้นตอนการค้นคืนแบบเดิม แสดงให้เห็นว่าขั้นตอนของงานวิจัยนี้สามารถค้นคืนเอกสารได้ตรงกับความต้องการของผู้ใช้มากกว่าวิธีขั้นตอนการค้นคืนแบบเดิม

โครงสร้างของออนโทโลยีเริ่มมีการพัฒนามากยิ่งขึ้น โดยเห็นได้จาก งานวิจัยของ (Mustafa *et al.*, 2008) ได้เสนอแบบจำลองของขั้นตอนการค้นคืนเอกสารแบบเชิงความหมาย โดยนำแนวคิดที่ว่าไม่มีคำค้นใดเพียงหนึ่งคำสามารถแทนบริบทของแนวความคิดนั้นๆ ได้ แต่ต้องมีหลายๆ คำในการแทนบริบทที่สนใจ ซึ่งงานวิจัยนี้ได้ปรับข้อมูลให้อยู่ในรูป RDF triples ที่มีลักษณะเป็นกราฟที่แสดงความหมาย 3 ประการ ได้แก่ ประธาน กริยา และคำในการทำนาย แสดงดังภาพที่ 8 ขั้นตอนในการขยายคำค้นได้ทำการเปรียบเทียบ RDF triples กับความสัมพันธ์ในออนโทโลยี คำที่ถูกใช้ในการขยายคำค้นคือคำที่มีความหมายเหมือนกัน และคำที่ใช้แทนคำหลักได้ เช่น คำย่อ โดยคำนวณน้ำหนักจากค่า *TFIDF* และค่าความคล้ายคลึง ข้อมูลที่ใช้ในการทดลองคือกลุ่มเอกสาร 300 เอกสาร โดยแต่ละเอกสารประกอบด้วย 30-50 RDF ผลการทดลองแสดงให้เห็นว่าการใช้ออนโทโลยีในการขยายคำค้นช่วยเพิ่มประสิทธิภาพของค่าความแม่นยำและค่าความระลึกได้ดีกว่าการค้นคืนแบบเดิม



ภาพที่ 8 กราฟ RDF triple ที่มีความสัมพันธ์โดยตรง

2. การค้นคืนเอกสารด้านกฎหมาย

การค้นคืนกฎหมายในปัจจุบันมีการใช้ฐานความรู้ที่ใช้แสดงถึงความรู้ด้านกฎหมายจากภายนอก เช่น พจนานุกรม และการแบ่งกลุ่ม ซึ่งไม่สามารถให้แนวความคิดที่เกี่ยวข้องกับกฎหมายได้อย่างถูกต้อง และยังใช้คนในการจัดเรียงลำดับของเอกสาร ถ้าเอกสารมีมากขึ้นทำให้เสียเวลาและค่าใช้จ่ายที่เพิ่มขึ้น งานวิจัยของ (Moens., 2001) ได้รวบรวมขั้นตอนที่ช่วยในการค้นคืนหลายขั้นตอน เช่น การสกัดคำ การแบ่งกลุ่ม การจำแนก การสรุปผล เป็นต้น

ออนโทโลยีทางกฎหมายมีการสร้างขึ้นเพื่อเก็บความรู้จากผู้เชี่ยวชาญให้อยู่ในรูปแบบฐานความรู้ งานวิจัยของ (Lame and Despres., 2005) ได้เสนอวิธีการสร้างออนโทโลยีและ การรวมออนโทโลยีจากออนโทโลยีที่มีอยู่ เนื่องจากออนโทโลยีด้านกฎหมาย จะมีการปรับปรุงเมื่อกฎหมายมีการเปลี่ยนแปลง และคำศัพท์ของกฎหมายเป็นทั้งคำทั่วไปและคำเฉพาะด้าน เช่น ด้านการแพทย์ ด้านวิทยาศาสตร์ เป็นต้น ดังนั้นงานวิจัยนี้จึงได้นำออนโทโลยีที่มีอยู่มารวมกันเพื่อปรับปรุงออนโทโลยีให้ตรงกับโครงสร้างของกฎหมายที่มีการเปลี่ยนแปลง การสร้างออนโทโลยีต้นแบบสร้างโดยใช้หลักของการประมวลผลภาษารธรรมชาติ โดยทำการสกัดคำที่ใช้เป็นโนดจากระบบวิเคราะห์หน้าที่ของคำ (Syntax) และทำการตัดคำที่ไม่ต้องการออก เช่น ตัวเลข คำเชื่อม เป็นต้น

การรวมออนโทโลยีใช้ออนโทโลยีที่ถูกสร้างขึ้นในแต่ละด้าน โดยตรวจสอบโครงสร้างที่คล้ายกันของแต่ละออนโทโลยี และเพิ่มโนดหรือความสัมพันธ์ให้กับออนโทโลยีใหม่ ข้อมูลที่ใช้ในการทดลองคือกฎหมายแพ่งและกฎหมายอาญา ของประเทศฝรั่งเศส ผลการทดลองแสดงให้เห็นว่าการปรับปรุงออนโทโลยีจากออนโทโลยีที่มีอยู่ ช่วยประหยัดเวลาและค่าใช้จ่ายในการสร้างออนโทโลยี และงานวิจัยของ (Kayed., 2005) ได้ทำการสร้างเหมือนข้อมูลกับกฎหมายเพื่อสร้างออนโทโลยีสำหรับกฎหมายอิเล็กทรอนิกส์ (e-laws) โดยออกแบบขั้นตอนวิธี 3 ขั้นตอนคือการสร้าง

ออนโทโลยี การจัดการออนโทโลยี และการดูแลออนโทโลยี ซึ่งทั้ง 3 ขั้นตอนทำงานแบบวนซ้ำ จนได้ออนโทโลยีที่ต้องการ รวมทั้งการทำการลดขนาดของ โหนดและความสัมพันธ์ให้มีขนาดเล็กลง

ออนโทโลยีสำหรับกฎหมายต่างประเทศมีการวิจัยเพิ่มมากขึ้นจนสามารถนำมาใช้ในงานวิจัยหลายๆ ด้านได้ เช่น งานวิจัยของ (Schweighofer and Geist., 2007) ได้ทำการปรับปรุงการค้นคืนกฎหมายของยุโรป โดยระบบเดิมได้ใช้การค้นคืนด้วยขั้นตอนการค้นคืนแบบบูลีน (Boolean search) โดยต้องทราบคำที่จะใช้ในการค้นหาให้ตรงกับความต้องการ เช่น คำที่มีความหมายเดียวกัน ซึ่งมีมากถึง 50,000 คำ ในเอกสารแต่ละกลุ่ม อีกทั้งคำที่ใช้ในกฎหมายต้องเป็นคำที่มีความยืดหยุ่น เพื่อสามารถปรับเข้าได้กับแต่ละสถานการณ์ที่เกิดขึ้น ดังนั้น คำเหล่านี้จึงมีความกำกวม ซึ่งสามารถใช้ได้ขึ้นอยู่กับความเป็นจริงในแต่ละคดีเท่านั้น และค่าความถี่ของคำไม่สามารถช่วยในการค้นคืนกฎหมายเหมือนเอกสารกลุ่มอื่นๆ อีกทั้งข้อความที่ตรงกับความต้องการมักปรากฏเพียง 1 ย่อหน้า หรือ 1 ประเด็น งานวิจัยนี้จึงได้สร้างต้นแบบขั้นตอนในการค้นคืนกฎหมายโดยใช้การขยายคำค้นด้วยออนโทโลยี และใช้คำค้นจากเอกสารที่ตรงกับความต้องการ ที่ผู้ใส่ส่งกลับคืนมา เมื่อนำ 2 ขั้นตอนมาช่วยในการขยายคำค้นจะได้กลุ่มของคำที่ใช้ในการเปรียบเทียบกับลำดับคำที่ตรงกับความต้องการจากฐานข้อมูล โดยได้ทดลองกับกฎหมายการเกษตร 1,770 ฉบับ ซึ่งให้ผลที่มีประสิทธิภาพมากกว่าระบบเดิม

งานวิจัยของ (Saravanan *et al.*, 2009) เสนอขั้นตอนวิธีการค้นคืนเอกสารเกี่ยวกับกฎหมาย เพื่อปรับปรุงประสิทธิภาพการค้นคืนกฎหมายในปัจจุบัน เนื่องจากเอกสารด้านกฎหมายเริ่มมีการจัดเก็บรูปแบบเอกสารอิเล็กทรอนิกส์มากขึ้น และระบบในปัจจุบันใช้คำค้นจากผู้ให้เพียงคำเดียวในการค้นคืนเอกสาร ทำให้ประสิทธิภาพในการค้นคืนคำ ไม่สามารถค้นคืนเอกสารที่ตรงกับความต้องการได้ งานวิจัยนี้จึงเสนอขั้นตอนวิธีค้นคืนโดยใช้ออนโทโลยีในการขยายคำค้น ซึ่งให้ผู้เชี่ยวชาญออกแบบโครงสร้างของออนโทโลยี โดยกลุ่มข้อมูลที่นำมาทดสอบคือกฎหมายการควบคุมการเช่า กฎหมายภาษี และกฎหมายพาณิชย์ โดยให้ผู้เชี่ยวชาญกฎหมายจำนวน 5 คน และบุคคลทั่วไปจำนวน 6 คน ทำการวัดประสิทธิภาพของค่าความแม่นยำ ค่าความระลึก และค่า F-measure เปรียบเทียบกับ Microsoft Search query ผลการทดลองแสดงให้เห็นว่าการขยายคำค้นโดยใช้ออนโทโลยีมีประสิทธิภาพในการค้นคืนเอกสารด้านกฎหมายได้ดีกว่าขั้นตอนการค้นคืนแบบเดิม ซึ่งคำค้นสามารถค้นคืนเอกสารทั่วไปเกี่ยวกับกฎหมายได้เท่านั้น แต่การขยายคำค้นโดยใช้ออนโทโลยีสามารถค้นคืนรายละเอียดของประเด็นต่างๆ ทางกฎหมายจากความสัมพันธ์ที่ผู้เชี่ยวชาญได้ออกแบบไว้

ตารางที่ 1 สรุปงานตามลำดับเวลางานวิจัย

ลำดับ	ปี	ชื่องานวิจัย	สิ่งที่เสนอ
1	2000	The Effectiveness of Information Retrieval with Local Context	- เสนอขั้นตอนวิธี Local Context Analysis - ค้นคืนเอกสารได้ ดีกว่าวิธี Local analysis และ Global analysis
2	2001	Innovative techniques for legal text retrieval	- แนวคิดในการค้นคืนกฎหมาย - การสกัดคำ การแบ่งกลุ่ม การจำแนก การสรุปผลช่วยในการค้นคืนกฎหมาย
3	2003	A Standard Framework for Personalization Via Ontology-Based Query Expansion	- รูปแบบการค้นคืนจากแนวความคิดของออนโทโลยี - เพิ่มคำค้นจากแนวคิดของออนโทโลยีทำให้ได้คำค้นที่ตรงกับความต้องการมากยิ่งขึ้น
4	2005	Updating Ontologies in the Legal Domain	- ขั้นตอนการสร้างออนโทโลยีและการรวมออนโทโลยี - ทำวิเคราะห์คำในโน้ตและความสัมพันธ์จากหลักของการประมวลผลธรรมชาติ
5	2005	Building e-Law Ontology: New Approach	- การใช้เหมืองข้อมูลในการสร้างออนโทโลยี - สามารถเพิ่มออนโทโลยีเกี่ยวกับกฎหมายอิเล็กทรอนิกส์
6	2007	Design and Implementation of Ontology-Based Query Expansion for Information Retrieval	- ใช้ออนโทโลยีในการขยายคำค้น - ค้นคืนเอกสารได้ ตรงความต้องการมากกว่าวิธีค้นคืนทั้งเอกสาร และค้นคืนโดยการขยายคำค้นแบบใช้ความถี่

ตารางที่ 1 (ต่อ)

ลำดับ	ปี	ชื่องานวิจัย	สิ่งที่เสนอ
7	2007	Legal Query Expansion using Ontologies and Relevance Feedback	- ใช้ออนโทโลยีค้นคืนเอกสารด้านกฎหมาย - ค้นคืนเอกสารโดยใช้ออนโทโลยี
8	2008	Ontology Based Semantic Information Retrieval	- ในแต่ละประโยคต้องมีคำหลายคำเป็นตัวแทนของประโยค โดยใช้ออนโทโลยีขยายคำ - สามารถค้นคืนเอกสารที่มีความเกี่ยวข้องกันออกมาได้มากกว่าการค้นคืนโดยใช้คำค้นเพียงคำเดียว
9	2008	Building an Ontology-Based Framework For Semantic Information Retrieval: Application To Breast Cancer	- ค้นคืนข้อมูลด้านการแพทย์ โดยใช้ออนโทโลยีมะเร็งเต้านม - ค้นคืนเอกสารที่สามารถป้อนคำค้นทั่วไปจากผู้ใช้ได้
10	2009	Improving legal information retrieval unsing an ontology framework	- ผลการค้นคืนสามารถค้นคืนเอกสารทางด้านกฎหมายได้ดีกว่าวิธีคำค้นหลักเพียงอย่างเดียว

อุปกรณ์และวิธีการ

อุปกรณ์

1. เครื่องคอมพิวเตอร์ หน่วยประมวลผล AMD Turion(tm) 64 Mobile Technology 2.0 GHz แรม 2 GB, หน่วยความจำรอง 100 GB
2. โปรแกรม Microsoft Visual Studio 2005
3. ระบบปฏิบัติการ Microsoft Windows Vista Home Premium

วิธีการ

งานวิจัยนี้ประกอบด้วยการทดลอง 3 การทดลอง โดยออนไลน์โทโลยีและขั้นตอนที่ใช้ในการทดลอง มีการปรับเปลี่ยนโครงสร้างความสัมพันธ์เพื่อให้สามารถค้นคืนคำพินาศาสาฎีกาให้มีประเด็นทางกฎหมายที่มีความหลากหลาย

1. ข้อมูลที่ใช้ในการทดลอง

ระบบสืบค้นคำพินาศาสาฎีกาและคำสั่งคำร้องสาฎีกา เริ่มใช้งานในปี พ.ศ. 2524 แสดงดังภาพที่ 9 ขั้นตอนในการสืบค้นคำพินาศาสาฎีกาใช้การสืบค้นโดยการค้นหาทั้งเอกสาร (Full text search) โดยรับคำค้นหลักจากผู้เข้ามาสืบค้นคำพินาศาสาฎีกาที่ได้เก็บรวบรวมในฐานข้อมูล และสามารถรับคำค้นได้ครั้งละหลายๆ คำค้น แต่ละคำค้นเชื่อมด้วย “และ”, “หรือ”, “ยกเว้น”, “ตามด้วย” ผลการสืบค้นคำพินาศาสาฎีกาในลักษณะนี้ ระบบจะไม่สนใจความหมายของคำ ทำให้ผลลัพธ์ของการสืบค้นคำพินาศาสาฎีกาได้คำพินาศาสาฎีกาที่ไม่แสดงความหลากหลายของประเด็นทางกฎหมาย

ข้อมูลที่ใช้ในการทดลองในงานวิจัย ผู้วิจัยได้รวบรวมคำพินาศาสาฎีกาตั้งแต่พ.ศ. 2542 ถึง พ.ศ. 2552 จากทั้งหมด 99,656 คำพินาศาสาฎีกาจากระบบสืบค้นคำพินาศาสาฎีกาและคำสั่งคำร้องสาฎีกา ประกอบด้วยคำพินาศาสาฎีกาคติมรดกและตัวเงิน โดยทำการแปลงข้อมูลและเก็บอยู่ในรูปของไฟล์อักษร (Text files) เพื่อสามารถประมวลผลในระบบที่ผู้วิจัยได้สร้างขึ้น

ภาพที่ 9 ระบบสืบค้นคำพิพากษาและคำสั่งคำร้องศาลฎีกา

2. วิธีวัดประสิทธิภาพ

การประเมินประสิทธิภาพของผู้เชี่ยวชาญ ประเมินจากผลลัพธ์ของการค้นคืนคำพิพากษาศาลฎีกาโดยพิจารณาจาก

1. คำพิพากษาศาลฎีกาที่ค้นคืนได้มีประเด็นทางกฎหมายที่ต้องวินิจฉัยสอดคล้องกับคำค้นหลักหรือไม่

2. การค้นคืนผลลัพธ์ ต้องมีประเด็นทางกฎหมายที่แตกต่างกันทุกๆ อันดับ

ประเด็นทางกฎหมายคือประเด็นต่างๆ ที่เกิดขึ้นในกฎหมาย เช่น ประเด็นเกี่ยวกับเช็ค ประเด็นเกี่ยวกับตั๋วแลกเงิน ประเด็นเกี่ยวกับตั๋วสัญญาใช้เงิน เป็นต้น

ข้อเท็จจริงที่แตกต่างกันคือการพิจารณาคำพิพากษาศาลฎีกาที่มีประเด็นทางกฎหมายเหมือนกัน แต่มีข้อเท็จจริงที่แตกต่างกัน เช่น ประเด็นทางกฎหมายคดี เช็ค จะมีข้อเท็จจริงที่แตกต่างกัน เช่น เช็คไม่มีการลงลายมือชื่อ เช็คที่ไม่มีการลงวันที่ เช็คที่มีการปลอมลายมือ เป็นต้น

ค่าความหลากหลายของประเด็นทางกฎหมายของคำพิพากษาศาลฎีกาคือค่าที่ใช้วัดประสิทธิภาพของงานวิจัยนี้ โดยพิจารณาจากกลุ่มคำค้นแต่กลุ่มที่สามารถค้นคืนคำพิพากษาศาลฎีกา ให้ค่าความหลากหลายของประเด็นทางกฎหมายมากขึ้นเพียงใด โดยค่าความหลากหลายของประเด็นทางกฎหมายมีค่าสูง แสดงว่าประเด็นและข้อเท็จจริงต่างๆ มีความหลากหลายทางกฎหมายมาก ซึ่งสามารถคำนวณค่าความหลากหลายของประเด็นทางกฎหมายได้จากสมการ (7)

3. ขั้นตอนวิธีที่ใช้ในการเปรียบเทียบ

3.1 ขั้นตอนวิธี Baseline

ขั้นตอนวิธี Baseline คือขั้นตอนที่ผู้วิจัยได้สร้างขึ้นในการจำลองระบบของการค้นคืนคำพิพากษาศาลฎีกา โดยรวบรวมคำพิพากษาศาลฎีกาจากระบบสืบค้นคำพิพากษาศาลฎีกา และมีขั้นตอนในการค้นคืนคำพิพากษาโดยใช้ค่าความถี่ของคำ *TFIDF* มีขั้นตอนการทำงานดังต่อไปนี้

1. รับคำค้นจากผู้ใช้
2. นำคำค้นจากผู้ใช้งานไปค้นคืนคำพิพากษาศาลฎีกา โดยหาคำนำน้าหนักของเอกสารจาก *TFIDF* ของคำค้นเพียงคำเดียว
3. เรียงลำดับของเอกสารจากมากไปน้อย
4. เลือกเอกสารตามจำนวนค่าหลากหลายที่ปรากฏในออนโทโลยี

3.2 ขั้นตอนวิธี Deka I

ขั้นตอนวิธี Deka I ใช้คำค้นเพียงคำเดียวที่ใช้ในการทดลองเข้าไปค้นคืนในระบบสืบค้นคำพิพากษาศาลฎีกา

3.3 ขั้นตอนวิธี Deka II

ขั้นตอนวิธี Deka II ใช้กลุ่มของคำค้นย่อยแต่ละกลุ่มที่ได้จากการขยายคำค้นโดยออนโทโลยีเข้าไปค้นคืนในระบบสืบค้นคำพิพากษาศาลฎีกา โดยเชื่อมคำด้วย “และ”

ขั้นตอนวิธี SCRO

ผู้วิจัยได้เสนอขั้นตอนการขยายคำค้นโดยใช้ความสัมพันธ์ของออนโทโลยี และนำกลุ่มคำค้นที่ได้ ช่วยในการค้นคืนคำพิพากษาศาลฎีกา ผู้วิจัยตั้งชื่อขั้นตอนนี้ว่า SCRO ย่อมาจาก Supreme Court sentence Retrieval unsig Ontology หรือการใช้ออนโทโลยีเพื่อการค้นคืนคำพิพากษาศาลฎีกา ซึ่งคือขั้นตอนที่เริ่มจากรับคำค้นจากผู้ใช้นำมาขยายคำโดยใช้ออนโทโลยี และนำกลุ่มคำค้นที่ได้ใช้ในการค้นคืนคำพิพากษาศาลฎีกา และทำการจัดลำดับของคำพิพากษาศาลฎีกาก่อนที่จะแสดงผลลัพธ์ให้กับผู้ใช้ ขั้นตอนหลักของ SCRO

งานวิจัยนี้ได้เสนอขั้นตอนวิธี SCRO ทั้งหมด 3 ขั้นตอนวิธี โดยแต่ละขั้นตอนวิธีได้ทำการปรับโครงสร้างและขั้นตอนวิธีเพื่อใช้ในการเปรียบเทียบประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมาย แต่ละขั้นตอนดังต่อไปนี้

1. ขั้นตอนวิธี SCRO

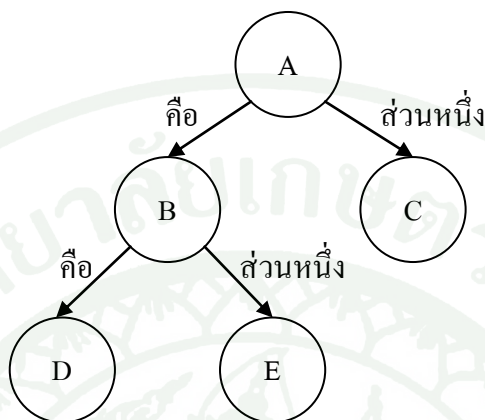
1.1 โครงสร้างออนโทโลยี

โครงสร้างของออนโทโลยีที่ใช้ในขั้นตอนวิธี SCRO เป็นโครงสร้างต้นไม้แสดงถึงความสัมพันธ์ระหว่างคำต่างๆ ประกอบด้วยโนดและความสัมพันธ์ที่เชื่อมระหว่างโนด 2 โนด ผู้เชี่ยวชาญเป็นผู้กำหนดคำในโนดโดยหาความสัมพันธ์แบบโครงสร้างต้นไม้คือมีโนดแม่และโนดลูก โดยโนดลูกที่เกิดขึ้นได้ โนดแม่ต้องเกิดก่อน

คำในโนดของออนโทโลยีสำหรับขั้นตอนวิธี SCRO ใช้คำนามหรือกริยาในการสร้างความสัมพันธ์ระหว่างคำ และความสัมพันธ์ที่ใช้เชื่อมระหว่างโนด เป็นความสัมพันธ์แบบ “คือ” (is-a) หรือ “ส่วนหนึ่งของ” (part-of) โดยโครงสร้างของออนโทโลยีแสดงดังภาพที่ 10

ขั้นตอนวิธีการขยายคำค้น ใช้ออนโทโลยีในการเพิ่มคำสำคัญโดยนำคำค้นที่ได้จากผู้ใช้เข้าไปค้นหาในออนโทโลยีและนำคำที่มีความสัมพันธ์กับคำค้นเพิ่มเข้ามาในชุดของคำค้น โดยค้นลงไปภายในออนโทโลยี 1 ระดับและเมื่อคำที่ถูกค้นขึ้นมามีความสัมพันธ์เชื่อม “คือ” ขั้นตอนวิธีจะทำการค้นหาลงไปอีก 1 ระดับ และหยุดเมื่อความสัมพันธ์ไม่ใช่ความสัมพันธ์ “คือ” ผลลัพธ์ของ

การขยายคำค้นจากขั้นตอนวิธี SCRO คำแต่ละคำที่ถูกขยายเพิ่มขึ้นจะสร้างเป็นกลุ่มของคำค้นเพียงกลุ่มเดียว



ภาพที่ 10 โครงสร้างของออนโทโลยีสำหรับขั้นตอนวิธี SCRO

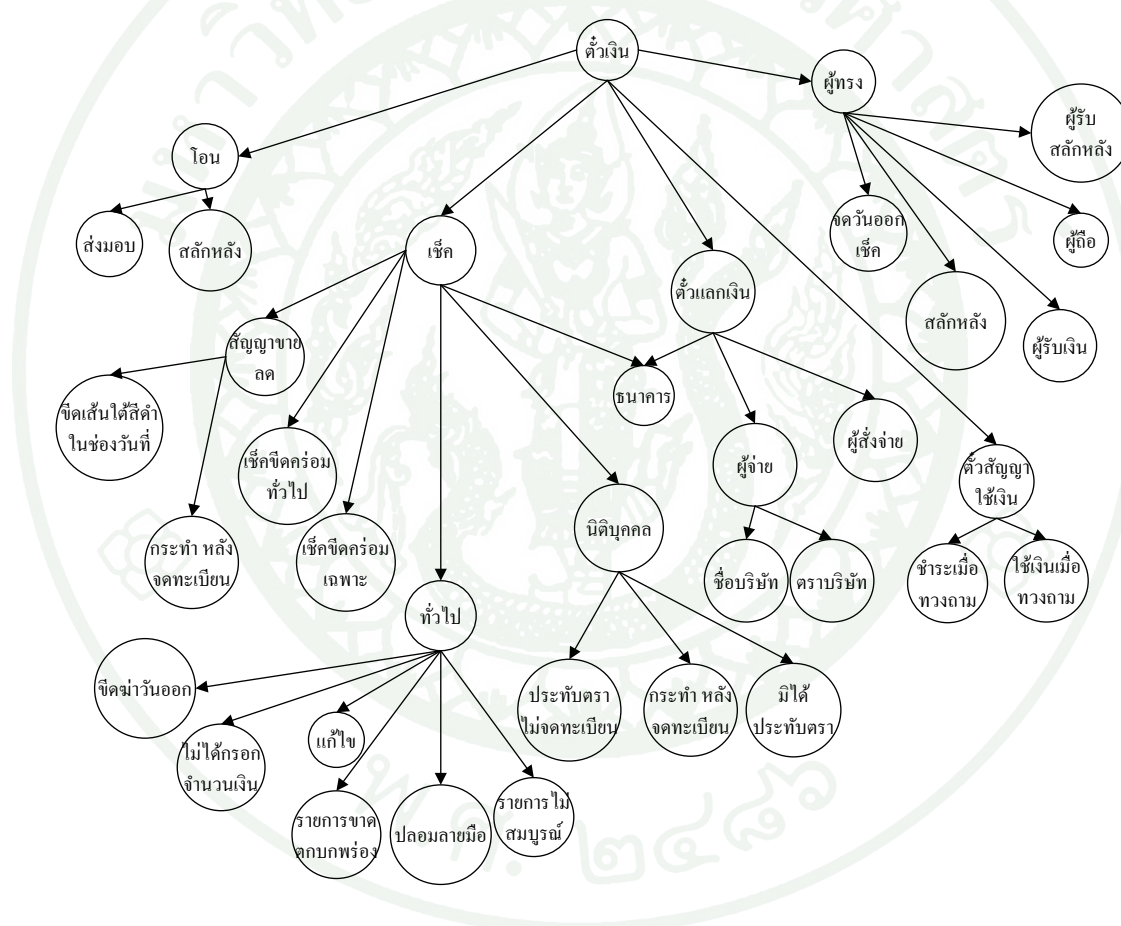
จากภาพที่ 10 โครงสร้างของออนโทโลยีประกอบด้วย โหนด A, B, C, D, E และความสัมพันธ์ที่เชื่อมระหว่างโหนดคือ “คือ” และ “ส่วนหนึ่ง” ถ้า A เป็นคำค้น กลุ่มคำในระดับแรกที่ถูกขยายคือ $\{B, C\}$ และ B สามารถขยาย $\{D\}$ เพิ่มเข้ามาในกลุ่มคำได้ ดังนั้นกลุ่มคำที่ขยายได้คือ $\{A, B, C, D\}$ ถ้า D คือคำค้น กลุ่มคำในระดับแรกที่ถูกขยายคือ $\{B\}$ และ B สามารถขยาย $\{A\}$ เพิ่มเข้ามาในกลุ่มคำค้นได้ ดังนั้นกลุ่มคำที่ขยายได้คือ $\{D, B, A\}$

1.2 ออนโทโลยีมรดก

โครงสร้างของออนโทโลยีมรดกสำหรับขั้นตอนวิธี SCRO ประกอบด้วยโหนด 55 โหนด และความสัมพันธ์ 56 ความสัมพันธ์ ซึ่งความสัมพันธ์ของออนโทโลยี “มรดก” มี 2 แบบคือคำที่มีความหมายเหมือนกันแทนด้วย “s” (synonym) และคำที่มีความสัมพันธ์กัน ดังภาพที่ 11

1.3 ออนโทโลยีตัวเงิน

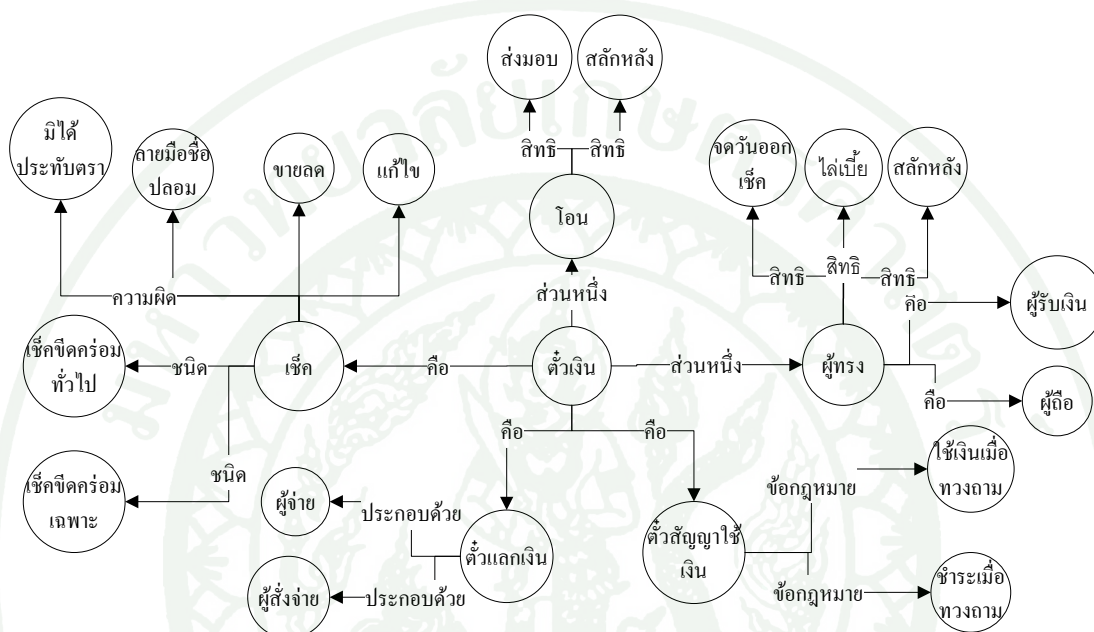
โครงสร้างของออนโทโลยีตัวเงินสำหรับขั้นตอนวิธี SCRO ประกอบด้วยโนด 36 โนด และความสัมพันธ์ 36 ความสัมพันธ์ มีการกำหนดค่าไว้ในส่วนของโนดเพียงอย่างเดียว โดยค่าในโนดเป็นได้ทั้งคำนามและกริยา ไม่มีการกำหนดค่าในส่วนของการความสัมพันธ์ ตัวอย่างค่าในโนด เช่น “ตัวเงิน” “แก้ไข” “ทั่วไป” เป็นต้น โครงสร้างความสัมพันธ์ของออนโทโลยี “ตัวเงิน” แสดงดังภาพที่ 12



ภาพที่ 12 ออนโทโลยีตัวเงินของขั้นตอนวิธี SCRO

จากภาพที่ 12 แสดงโครงสร้างของความสัมพันธ์ระหว่างคำที่เกิดขึ้นเฉพาะคดีที่เกี่ยวข้องกับกฎหมายคดีตัวเงิน ตัวอย่างของความสัมพันธ์ของออนโทโลยี เช่น “ตัวเงิน” มีความสัมพันธ์กับ “โอน” “เช็ค” “ตัวแลกเงิน” “ผู้ทรง” เป็นต้น

ในการทดลองใช้ความสัมพันธ์ในออนโทโลยีในการคำนวณค่าสัมประสิทธิ์ของแต่ละคำ ผู้เชี่ยวชาญจึงปรับโครงสร้างของออนโทโลยีให้มีคำที่ใช้ออกความสัมพันธ์พื้นฐานของออนโทโลยี เช่น “คือ” “ประกอบด้วย” “ส่วนหนึ่ง” เป็นต้น แสดงดังภาพที่ 13

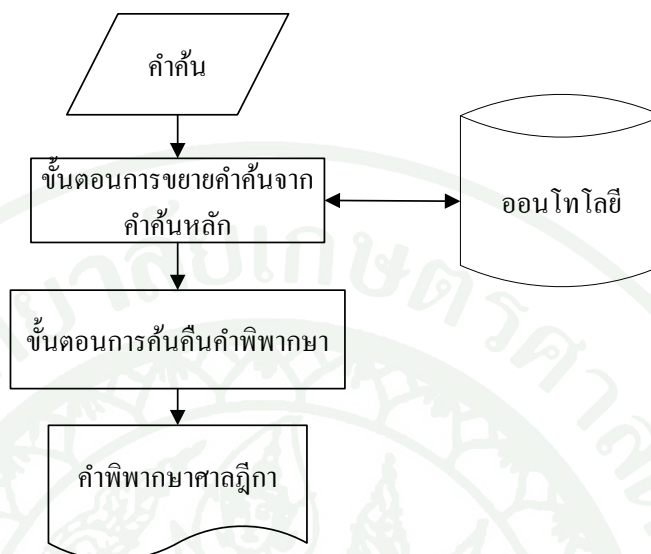


ภาพที่ 13 ออนโทโลยีตัวเงินสำหรับการทดลองที่ 3

ภาพที่ 13 แสดงโครงสร้างออนโทโลยีตัวเงิน ตัวอย่างความสัมพันธ์ เช่น “ตัวเงิน คือ เช็ค” “ผู้ทรง คือ ผู้รับเงิน”

จากภาพที่ 13 ตัวอย่างของการแทนค่าในโนด คำในโนดใช้แทนคำนามหรือกริยา ส่วนคำในความสัมพันธ์เป็นคำที่บอกถึงความสัมพันธ์ระหว่างโนดเท่านั้น เช่น คำเหมือน คำที่เป็นส่วนประกอบ เป็นต้น โดยไม่ถูกเพิ่มเข้าไปในกลุ่มของคำค้นใหม่ โครงสร้างความสัมพันธ์ของออนโทโลยีที่ปรากฏ เช่น โนดแม่คือ “ตัวเงิน” โนดลูกที่เกิดขึ้นได้ หรือคำที่เกี่ยวข้องคือ “ผู้ทรง” “โอน” “เช็ค” “ตัวเงิน” และ “ตัวสัญญาใช้เงิน”

1.4 ขั้นตอนวิธี SCRO



ภาพที่ 14 ขั้นตอนการทำงานของขั้นตอนวิธี SCRO

จากภาพที่ 14 แสดงขั้นตอนวิธี SCRO ในการค้นคืนคำพิพาทศาลฎีกา เริ่มจากผู้ใช้อัปโหลดคำค้นเข้ามายังระบบ จากนั้นขั้นตอนจะนำคำค้น ทำการค้นหาความสัมพันธ์ในออนโทโลยีที่ผู้เชี่ยวชาญได้ออกแบบไว้ ผลลัพธ์ที่ได้เป็นกลุ่มของคำสำคัญที่ได้จากออนโทโลยี และนำกลุ่มของความสัมพันธ์เข้าไปค้นคืนคำพิพาทศาลฎีกา ซึ่งมีรายละเอียดดังต่อไปนี้

- 1) คำค้นคือคำที่ใช้ในการค้นคืนคำพิพาทศาลฎีกา
- 2) ขั้นตอนการขยายคำค้นคือขั้นตอนการเพิ่มคำค้นให้กับคำค้นที่รับมาจากผู้ใช้งาน โดยอาศัยความสัมพันธ์ของออนโทโลยี ขั้นตอนวิธีจะดึงคำที่มีความสัมพันธ์กับคำค้นหลักออกมาในทุกๆ ความสัมพันธ์ สร้างเป็นกลุ่มของคำค้นใหม่ ความสัมพันธ์ของคำแบ่งเป็น 2 กรณีตามโครงสร้างของออนโทโลยี ดังนี้

2.1) คำที่มีความหมายเหมือนกันคือคำที่เขียนแตกต่างกันแต่มีความหมายเหมือนกัน ที่สามารถนำมาใช้แทนกันได้ เช่น คำว่า “ตาย” มีความหมายเหมือนคำว่า “มรณภาพ”, “ความตาย”, “ถึงแก่กรรม” เป็นต้น

2.2) คำที่มีความสัมพันธ์กันคือคำที่มีความหมายไม่เหมือนกัน แต่มีความสัมพันธ์กัน เมื่อมีคำใดเกิดขึ้น จะมีอีกคำเกิดขึ้นด้วยเสมอ เช่น “ผู้ตาย” มีความสัมพันธ์กับคำว่า “มรดก” และคำว่า “ทายาท” โดยความสัมพันธ์ของแต่ละคำถูกกำหนดไว้ในออนโทโลยีที่สร้างจากผู้เชี่ยวชาญ

3) ขั้นตอนการค้นคืนคำพินาศศาสตร์ ขั้นตอนนี้ใช้เวกเตอร์สเปซโมเดลคำนวณน้ำหนักของคำพินาศศาสตร์ โดยเลือกเอกสารจากกลุ่มของคำค้นใหม่ ซึ่งคำนวณน้ำหนักของเอกสาร และเลือกคำพินาศศาสตร์ที่มีค่าน้ำหนักมากที่สุดออกมาตามลำดับ ดังสมการ (8)

$$D_i = \sum (w_{i,j} * k_j) \quad (8)$$

โดย	D_i	คือ ค่าน้ำหนักของเอกสาร
	k_j	คือ ค่าสัมประสิทธิ์ของคำ
	w_i	คือ น้ำหนักของคำ
	i	คือ ลำดับเอกสาร
	j	คือ ลำดับของคำ

4) คำพินาศศาสตร์คือผลลัพธ์ที่ได้จากขั้นตอนวิธี SCRO โดยเรียงลำดับผลลัพธ์ตามน้ำหนักของคำพินาศศาสตร์จากมากไปน้อย

ค่าสัมประสิทธิ์ของคำ พิจารณาจากชนิดของความสัมพันธ์ที่เชื่อมระหว่างโนด โดยค่าสัมประสิทธิ์ของคำเริ่มต้นมีค่าเท่ากับ 1 และลดลงตามชนิดของความสัมพันธ์ ค่าสัมประสิทธิ์ที่มากแสดงว่าคำนั้นมีความสำคัญมาก ในทางกลับกันคำที่มีค่าสัมประสิทธิ์น้อย แสดงว่ามีความสำคัญน้อย ค่าสัมประสิทธิ์แสดงดังตารางที่ 2

ตารางที่ 2 ค่าสัมประสิทธิ์ของความสัมพันธ์

ความสัมพันธ์	สัมประสิทธิ์ความสำคัญของโนด	
	โนดลูก	โนดแม่
คือ	ค่าสัมประสิทธิ์ลดลง $\frac{1}{4}$ ของค่า สัมประสิทธิ์ปัจจุบัน	ค่าสัมประสิทธิ์ลดลง $\frac{1}{4}$ ของค่า สัมประสิทธิ์ปัจจุบัน
อื่นๆ	ค่าสัมประสิทธิ์ลดลง $\frac{1}{2}$ ของค่า สัมประสิทธิ์ปัจจุบัน	ค่าสัมประสิทธิ์ลดลง $\frac{3}{4}$ ของค่า สัมประสิทธิ์ปัจจุบัน

จากภาพที่ 2 ตัวอย่างของการคำนวณค่าสัมประสิทธิ์ความสำคัญของค่า จากตารางที่ 1 ถ้าโนด “ตัวเงิน” เป็นคำค้นหลัก ซึ่งมีค่าสัมประสิทธิ์ความสัมพันธ์คือ 1 ดังนั้น โหนด “เช็ค” เป็นโนดลูก มีความสัมพันธ์ “คือ” มีค่าสัมประสิทธิ์ความสัมพันธ์เท่ากับ 0.75 และ โหนด “โอน” เป็นโนดลูก มีความสัมพันธ์ “ส่วนหนึ่ง” มีค่าสัมประสิทธิ์ความสัมพันธ์เท่ากับ 0.5 เป็นต้น

2. ขั้นตอนวิธี SCRO I

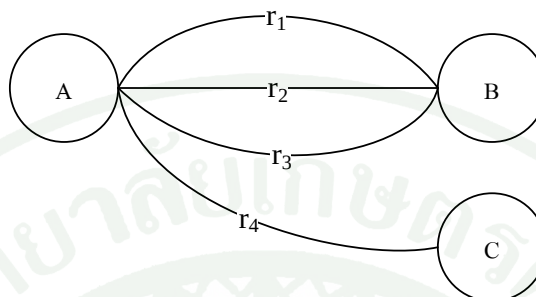
2.1 โครงสร้างออนโทโลยี

โครงสร้างของออนโทโลยีที่ใช้ในขั้นตอนวิธี SCRO I เป็นโครงสร้างออนโทโลยีที่ประกอบด้วยโนดและความความสัมพันธ์ที่ไม่มีทิศทาง ผู้เชี่ยวชาญเป็นผู้กำหนดค่าในโนดและความสัมพันธ์ ตามประเด็นทางกฎหมายที่เกิดขึ้น โดยโนด 1 โหนดสามารถมีได้หลายความสัมพันธ์ แต่ความสัมพันธ์ 1 ความสัมพันธ์ต้องเชื่อมระหว่างโนด 2 โหนดเท่านั้นแสดงดังภาพที่ 15

คำในโนดและความสัมพันธ์ของออนโทโลยีสำหรับขั้นตอนวิธี SCRO I ใช้คำนามหรือกริยา ในการสร้างความสัมพันธ์ระหว่างคำ โดยความสัมพันธ์แสดงถึงประเด็นทางกฎหมายที่มีความหลากหลาย

ขั้นตอนวิธี SCRO I ทำการขยายคำค้นโดยใช้ความสัมพันธ์ของออนโทโลยี ออกเป็นกลุ่มย่อยๆ ประกอบด้วยโนด 2 โหนดและความสัมพันธ์ 1 ความสัมพันธ์ที่สัมพันธ์กับโนดทั้ง 2 ซึ่ง

ผลลัพธ์ของการขยายคำค้นจากขั้นตอนวิธี SCRO I จะสร้างกลุ่มคำค้นย่อยๆ ตามเส้นความสัมพันธ์ของออนโทโลยี



ภาพที่ 15 โครงสร้างออนโทโลยีสำหรับขั้นตอนวิธี SCRO I

จากภาพที่ 15 โครงสร้างของออนโทโลยีประกอบด้วย โหนด เช่น A, B และ C และความสัมพันธ์คือ r_1, r_2, r_3 และ r_4 ความสัมพันธ์ระหว่างโหนด A และ B คือ $\{r_1, r_2, r_3\}$ และความสัมพันธ์ระหว่าง A และ C คือ $\{r_4\}$

ตัวอย่างการขยายคำค้นจากภาพที่ 19 ถ้า A เป็นคำค้น กลุ่มคำที่ขยายได้คือ $\{A, r_1, B\}$, $\{A, r_2, B\}$, $\{A, r_3, B\}$ และ $\{A, r_4, C\}$

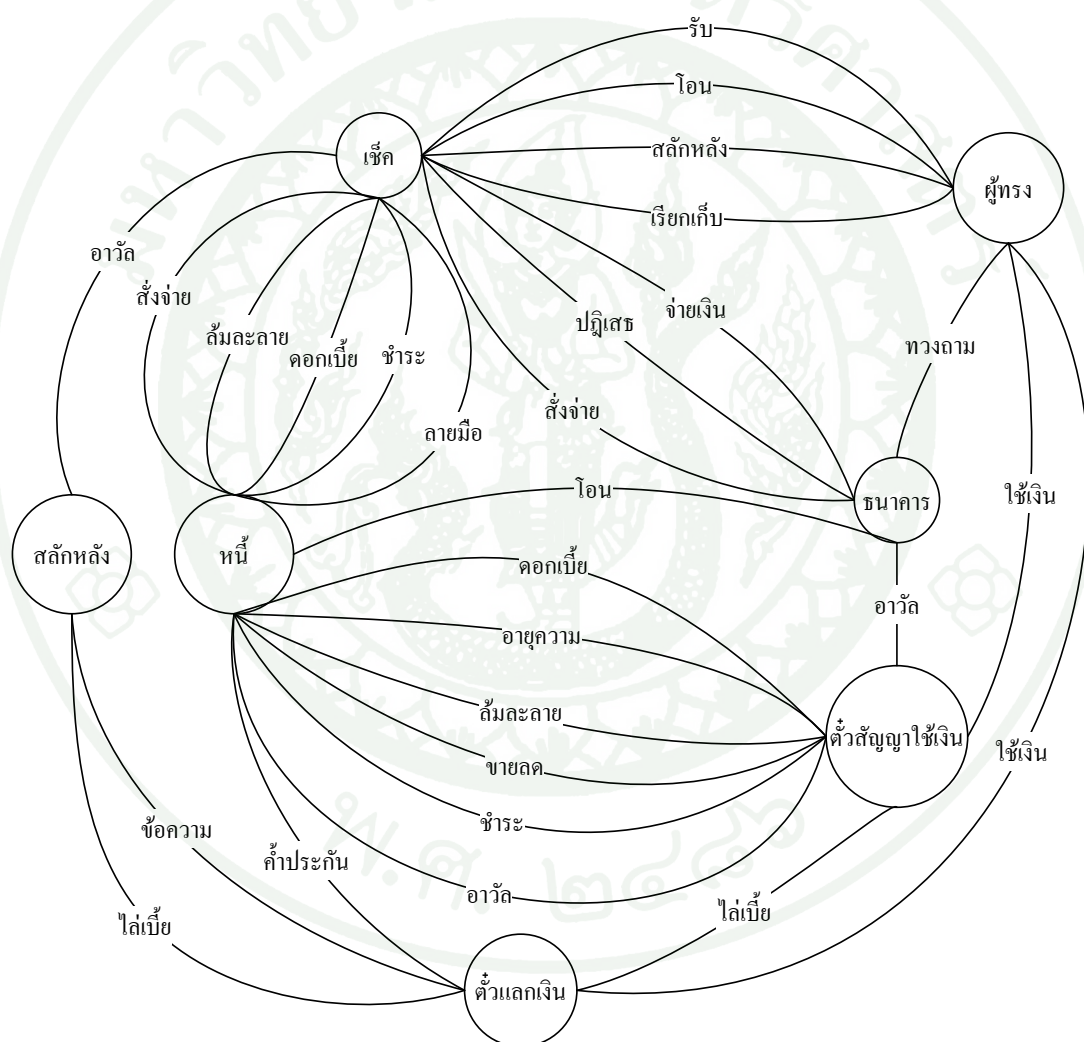
กลุ่มคำค้นที่ได้จากการขยายความสัมพันธ์ของโครงสร้างออนโทโลยีลักษณะนี้ ทำให้ได้กลุ่มคำค้นที่แต่ละกลุ่มแสดงถึงประเด็นต่างๆ เนื่องจากในการสร้างออนโทโลยี ความสัมพันธ์แสดงถึงประเด็นต่างๆ ของกฎหมายระหว่างโหนด 2 โหนดที่ใช้แทนลักษณะของแต่ละประเด็นที่ได้จากผู้เชี่ยวชาญ เมื่อนำกลุ่มคำค้นย่อย แต่ละกลุ่มแยกกันค้นคืนคำพิพากษาศาลฎีกา ทำให้แต่ละกลุ่มสามารถค้นคืนประเด็นที่ตรงกันกลุ่มๆ นั้นได้ ดังนั้น เมื่อเราพิจารณาผลลัพธ์ลำดับที่ 1 ของแต่ละกลุ่ม จะทำให้ได้คำพิพากษาศาลฎีกาที่มีความหลากหลายทางประเด็นของกฎหมาย

2.2 ออนโทโลยีมรดก

โครงสร้างของออนโทโลยีมรดกสำหรับขั้นตอนวิธี SCRO I สร้างจากคำพิพากษาศาลฎีกา และประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 6 มรดก มาตรา 1599 ถึงมาตรา 1754 จำนวน

2.3 ออนโทโลยีตัวเงิน

โครงสร้างของออนโทโลยีตัวเงินสำหรับขั้นตอนวิธี SCRO I สร้างจากคำพิพาทกษาศาสดาฎีกา และประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน มาตรา 898 ถึงมาตรา 1001 จำนวน 200 คำพิพาทกษา ออนโทโลยีประกอบด้วยโนด 7 โนดและความสัมพันธ์ 28 ความสัมพันธ์ ดังภาพที่ 17

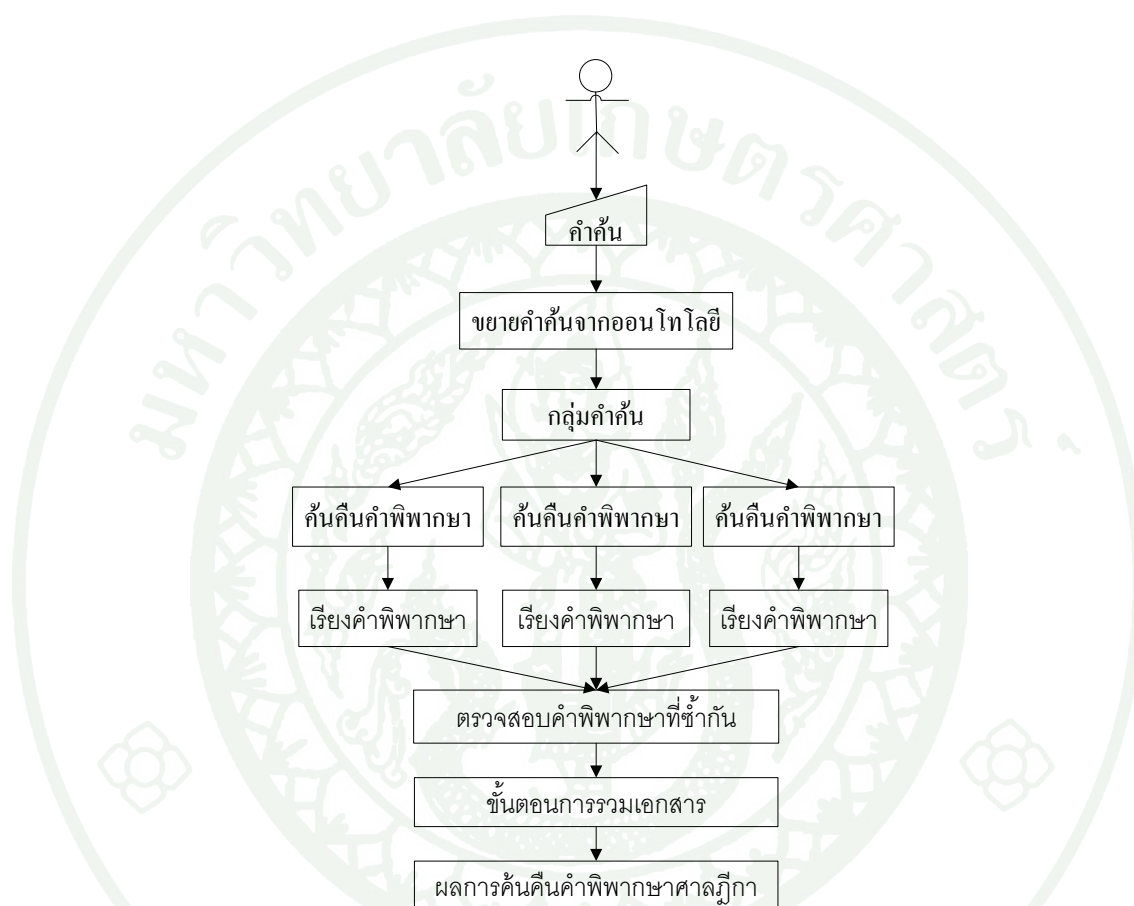


ภาพที่ 17 ออนโทโลยีตัวเงินสำหรับขั้นตอนวิธี SCRO I

จากภาพที่ 17 ความสัมพันธ์ของออนโทโลยีแสดงถึงประเด็นต่างๆ ระหว่างโนด 2 โนด เมื่อนำคำมาขยายในออนโทโลยี จะได้กลุ่มคำที่แสดงถึงความหลากหลายของประเด็นทางกฎหมาย

เช่น ความสัมพันธ์ระหว่างโนด “เชื้อ” และโนด “หนึ” คือ “ส่งจ่าย” “ล้มละลาย” “ดอกเบี้ย” “ชำระ” “ลายมือ” ตามลำดับ

2.4 ขั้นตอนวิธี SCRO I



ภาพที่ 18 แผนภาพขั้นตอนวิธี SCRO I และขั้นตอนวิธี SCRO II

จากภาพที่ 18 แผนภาพขั้นตอนวิธี SCRO I และ SCRO II มีภาพรวมของการทำงานเหมือนกันคือรับคำค้นจากผู้ใช้งาน ขยายคำค้นโดยใช้ออนไลน์ให้ได้กลุ่มคำค้นย่อย และนำไปค้นคืนคำพิพากษาศาลฎีกา โดยเลือกคำพิพากษาลำดับที่ 1 ของแต่ละกลุ่มคำค้นย่อยไม่ให้ซ้ำกัน ซึ่งแสดงแต่ละขั้นตอนโดยละเอียดดังต่อไปนี้

ตารางที่ 3 ขั้นตอนวิธี SCRO I

Algorithm: SCRO I

Input: Ontology, keyword

Begin

Word = {}

Query = {}

Word = OntologyTraversal (keyword)

Query = QueryGenerate(Word, $\emptyset$)

For each q_i in Query

D_i = DocumentRetrieval (q_i)

Ranking(D_i)

End For

CombineResults = checkDuplicate(D)

End

Output: CombineResults

จากตารางที่ 3 ข้อมูลเข้าคือคำค้นและออนโทโลยีในการทำงาน ขั้นตอนวิธี SCRO I จะมีการเก็บกลุ่มของคำค้นย่อย จากโครงสร้างของออนโทโลยี จากนั้นนำคำค้นแต่ละกลุ่มไปค้นคืนคำพิพาทศาสตร์ฎีกา และเลือกเอกสารลำดับที่ 1 ของแต่ละกลุ่มคำค้นย่อย โดยตรวจสอบไม่ให้มีหมายเลขคำพิพาทศาสตร์ฎีกาเดียวกัน ซึ่งมีขั้นตอนโดยละเอียดดังต่อไปนี้

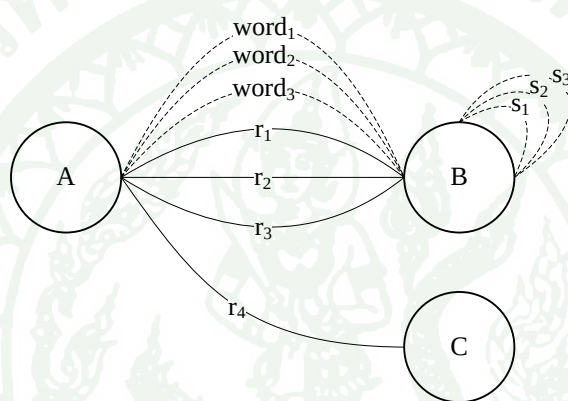
1) คำค้นคือคำที่ใช้ในการค้นคืนคำพิพาทศาสตร์ฎีกา

2) การขยายคำค้นจากออนโทโลยีคือขั้นตอนการขยายคำค้นจากโครงสร้างความสัมพันธ์ของออนโทโลยี โดยขั้นตอนจะนำคำค้น ไปเปรียบเทียบแต่ละ โหนดและแต่ละ ความสัมพันธ์ของออนโทโลยี เพื่อหาคำที่ตรงกับคำค้น โดยขั้นตอนการขยายคำแบ่งเป็น 2 กรณี

- เมื่อคำค้นตรงกับ โหนด จะขยายความสัมพันธ์ที่เชื่อมโยงกับ โหนดนั้น และ โหนดที่มีความสัมพันธ์ด้วยทั้งหมด สร้างเป็นกลุ่มของคำค้นใหม่ ตัวอย่างเช่น จากภาพที่ 15 ถ้า A เป็นคำค้นกลุ่มของคำค้นที่ขยายได้คือ $\{A, B, C, r_1, r_2, r_3, r_4\}$

- ถ้าคำค้นตรงกับความสัมพันธ์ จะขยายโนดทั้ง 2 ที่เชื่อมความสัมพันธ์นั้นออกมา ทั้งหมดสร้างเป็นกลุ่มของคำค้นใหม่ ตัวอย่างเช่น จากภาพที่ 15 ถ้า r_1 เป็นคำค้น กลุ่มของคำค้นที่ขยายได้คือ $\{A, B, r_1\}$

3) กลุ่มคำค้นคือกลุ่มของคำค้นที่ได้จากขั้นตอนการขยายคำค้น โดยโครงสร้างของกลุ่มของคำแต่ละกลุ่มประกอบด้วยกลุ่มของคำ 3 คำคือ โหนดที่สัมพันธ์กับคำหลัก ความสัมพันธ์ที่เชื่อมระหว่างโนดและคำที่มีความสัมพันธ์กับคำหลัก



ภาพที่ 19 โครงสร้างออนโทโลยี SCRO I

จากภาพที่ 19 เมื่อพิจารณาโครงสร้างของออนโทโลยี ถ้า A เป็นคำค้น กลุ่มคำค้นขยายได้ คือ $\{A, r_1, B\}$, $\{A, r_2, B\}$, $\{A, r_3, B\}$, $\{A, word_1, B\}$, $\{A, word_2, B\}$, $\{A, word_3, B\}$ และ $\{A, r_4, C\}$ ตามลำดับ

สำหรับขั้นตอนวิธี SCRO II โครงสร้างของออนโทโลยีจะถูกปรับเปลี่ยน โดยส่วนของ *word* และ *s* เป็นส่วนของออนโทโลยีระดับที่ 2 เนื่องจากโนดแทนคำนามและความสัมพันธ์แทนคำกริยา

3.4 ขั้นตอนค้นคืนคำพิพาทคือการนำกลุ่มของคำค้นทั้งหมด ค้นคืนคำพิพาทศาลฎีกาโดยหาน้ำหนักจากค่า *TFIDF* จากสมการ (2)

3.6 ขั้นตอนการเรียงคำพินาศคือการเรียงลำดับของคำพินาศศาสตร์ที่ค้นคืนออกมาได้ โดยทำการเรียงน้ำหนักของเอกสารจากมากไปน้อย

3.7 ขั้นตอนการตรวจสอบคำพินาศที่ซ้ำกันคือการตรวจสอบคำพินาศลำดับที่ 1 ของแต่ละกลุ่มไม่ให้ได้คำพินาศศาสตร์ที่หมายเลขพินาศเหมือนกัน โดยถ้าหมายเลขคำพินาศเหมือนกัน จะเลือกคำพินาศลำดับถัดไปจากคำพินาศที่มีน้ำหนักน้อยกว่า

3.8 ขั้นตอนการรวมเอกสารคือการรวมผลลัพธ์ของทุกๆ คำพินาศศาสตร์ที่ได้ทำการตรวจสอบไม่ให้ซ้ำกันแล้ว แสดงผลลัพธ์ให้กับผู้ใช้

3. ขั้นตอนวิธี SCRO II

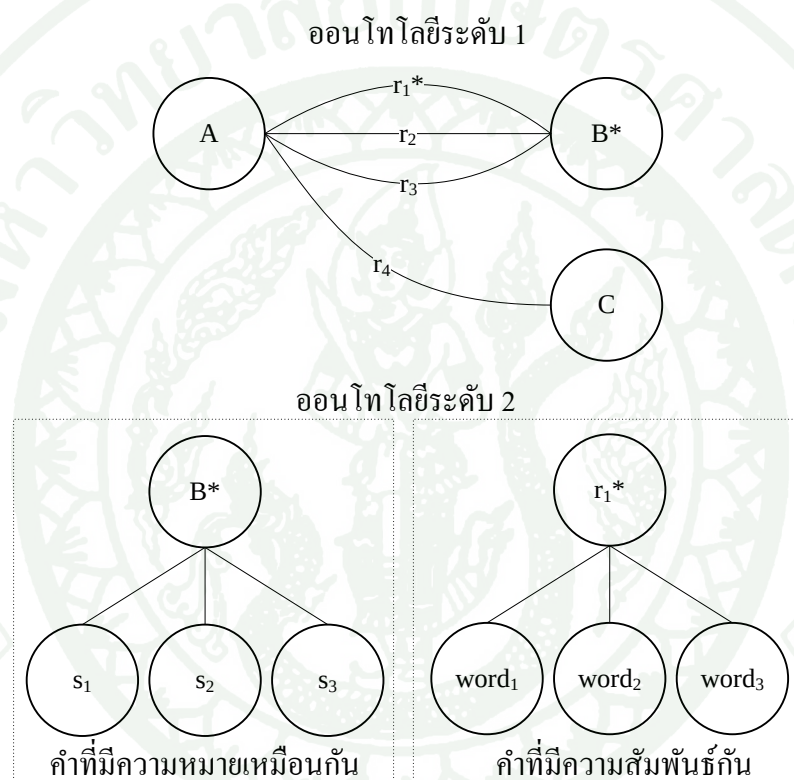
3.1 โครงสร้างออนโทโลยี

โครงสร้างของออนโทโลยีที่ใช้ในขั้นตอนวิธี SCRO II เป็นโครงสร้างออนโทโลยีระดับ 2 ชั้นคือออนโทโลยีหลัก หรือออนโทโลยีระดับที่ 1 และออนโทโลยีระดับชั้น หรือออนโทโลยีระดับที่ 2 โดยออนโทโลยีหลัก โดยชั้นออนโทโลยีระดับที่ 1 ประกอบด้วยโนดและความสัมพันธ์ที่ไม่มีทิศทาง ผู้เชี่ยวชาญเป็นผู้กำหนดคำในโนดและความสัมพันธ์ ตามประเด็นทางกฎหมายที่เกิดขึ้น โดยโนด 1 โนดสามารถมีได้หลายความสัมพันธ์ แต่ความสัมพันธ์ 1 ความสัมพันธ์ต้องเชื่อมระหว่างโนด 2 โนดเท่านั้น โดยแสดงถึงโครงสร้างความสัมพันธ์พื้นฐานของประเด็นต่างทางกฎหมายที่หลากหลาย

ออนโทโลยีชั้นที่ 2 คือส่วนขยายของออนโทโลยีหลัก แบ่งเป็น 2 ชนิดคือคำที่ขยายจากโนด โดยขยายคำที่มีความหมายเหมือนกัน และคำที่ขยายจากความสัมพันธ์ จะขยายคำที่มีความสัมพันธ์กัน คำในโนดของออนโทโลยีจะเป็นคำนามเท่านั้น และคำในความสัมพันธ์จะเป็นคำกริยา แสดงดังภาพที่ 22

ออนโทโลยีสำหรับขั้นตอนวิธี SCRO II คำในโนดของออนโทโลยีเป็นคำนามเท่านั้น และคำในความสัมพันธ์ของออนโทโลยีเป็นคำกริยาเท่านั้น

ขั้นตอนวิธี SCRO II ทำการขยายคำค้นโดยใช้ความสัมพันธ์ของออนโทโลยี ออกเป็นกลุ่มย่อยๆ ประกอบด้วยโนด 2 โนดและความสัมพันธ์ 1 ความสัมพันธ์ที่สัมพันธ์กับโนดทั้ง 2 ซึ่งผลลัพธ์ของการขยายคำค้นจากขั้นตอนวิธี SCRO I จะสร้างกลุ่มคำค้นย่อยๆ ตามเส้นความสัมพันธ์ของออนโทโลยี และสามารถขยายไปในออนโทโลยีระดับที่ 2 ได้เมื่อมี “*” ซึ่งเป็นส่วนกำหนดการขยาย



ภาพที่ 20 โครงสร้างออนโทโลยีสำหรับขั้นตอนวิธี SCRO II

จากภาพที่ 20 แสดงโครงสร้างของออนโทโลยี 2 ระดับ โดยระดับ 1 เป็นโครงสร้างของออนโทโลยีทั่วไป ประกอบด้วย โนดและความสัมพันธ์ ซึ่งในแต่ละโนดหรือความสัมพันธ์จะมี * เพื่อเชื่อมไปยังออนโทโลยีในชั้นที่ 2 โดยในชั้นนี้ ความสัมพันธ์หลักจะถูกขยายจากคำที่มีความสัมพันธ์กับคำหลัก ส่วนโนดจะถูกขยายจากคำที่มีความหมายเหมือนกัน

ขั้นตอนในการขยายคำค้นจะทำการขยายคำค้น โดยแบ่งออกเป็นกลุ่มย่อยๆ ตามความสัมพันธ์ของออนโทโลยี โดยแต่ละกลุ่มประกอบด้วยโนดที่มีความสัมพันธ์กัน 2 โนดและ

ความสัมพันธ์ที่เชื่อมระหว่างโนด ตัวอย่างของการขยายคำค้น เช่น คำค้น A กลุ่มคำค้นที่ขยายได้คือ $\{A, r_1, B\}$, $\{A, word_1, B\}$, $\{A, word_2, B\}$, $\{A, word_3, B\}$, $\{A, r_2, B\}$, $\{A, r_3, B\}$ และ $\{A, r_4, C\}$ ถ้า B เป็นคำค้น ได้กลุ่มคำคือ $\{A, B, r_1\}$, $\{A, B, r_2\}$, $\{A, B, r_3\}$, $\{A, B, s_1\}$, $\{A, B, s_2\}$ และถ้าคำค้น C กลุ่มของคำค้นที่ขยายได้คือ $\{A, r_4, C\}$

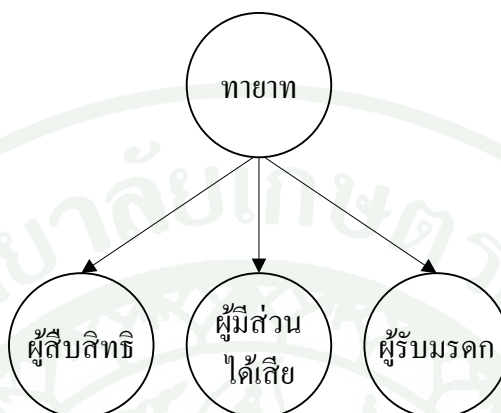
ถ้าโนดที่เป็นคำค้นมี * สามารถทำการขยายคำค้นไปในออนโทโลยีระดับที่ 2 ได้ ตัวอย่างเช่น คำค้น B กลุ่มคำค้นที่ขยายได้คือ $\{A, r_1, B\}$, $\{A, r_2, B\}$, $\{A, r_3, B\}$, $\{A, r_1, S_1\}$, $\{A, r_2, S_1\}$, $\{A, r_3, S_1\}$, $\{A, r_1, S_2\}$, $\{A, r_2, S_2\}$, $\{A, r_3, S_2\}$, $\{A, r_1, S_3\}$, $\{A, r_2, S_3\}$, $\{A, r_3, S_3\}$ คำค้น r_1 กลุ่มของคำค้นที่ขยายได้คือ $\{A, r_1, B\}$, $\{A, word_1, B\}$, $\{A, word_2, B\}$, $\{A, word_3, B\}$ คำค้น r_2 กลุ่มของคำค้นที่ขยายได้คือ $\{A, B, r_2\}$

ผู้เชี่ยวชาญเป็นผู้กำหนดความสัมพันธ์ของคำระหว่างโนดและความสัมพันธ์ และนำกลุ่มความสัมพันธ์ที่ได้ มาสร้างเป็นออนโทโลยี โดยโนดที่มีความหมายเหมือนกันจะนำคำที่เป็นพื้นฐานสร้างไว้ในออนโทโลยีระดับที่ 1 และคำที่มีความสัมพันธ์กันใช้เป็นคำขยายของโนด โดยสร้างไว้ในออนโทโลยีระดับที่ 2 ถ้าความสัมพันธ์ที่เชื่อมโยงระหว่างโนด เป็นคำที่มีความหมายเหมือนกัน และทั้ง 2 โหนดเป็นคำเดียวกัน จะนำคำของความสัมพันธ์ที่เป็นคำพื้นฐานมากกว่า มาใช้เป็นความสัมพันธ์ของออนโทโลยีระดับที่ 1 และคำของความสัมพันธ์ที่เหลือจะใช้ขยายความสัมพันธ์หลักของออนโทโลยีระดับที่ 1 ซึ่งนำมาสร้างไว้ในออนโทโลยีระดับที่ 2 และความสัมพันธ์ที่นอกเหนือจากนี้จะอยู่ในออนโทโลยีระดับที่ 1 ทั้งหมด

ผลของการขยายคำค้นในขั้นตอนวิธี SCRO II กลุ่มของคำค้นแบ่งออกเป็นกลุ่มย่อยๆ ตามประเด็นต่างๆ ของกฎหมาย ที่ปรากฏในความสัมพันธ์ของออนโทโลยี เช่น r_1 สามารถขยายได้ดังนี้ $\{A, r_1, B\}$, $\{A, word_1, B\}$, $\{A, word_2, B\}$, $\{A, word_3, B\}$ แต่ถ้าเป็นขั้นตอนวิธี SCRO I ในการทดลองที่ 2 คำค้น r_1 ขยายได้เพียง $\{A, r_1, B\}$ เท่านั้น

ถ้ากลุ่มคำค้นที่ใช้เป็นกลุ่มเดียวคือ $\{A, r_1, word_1, word_2, word_3, B\}$ เหมือนการขยายคำค้นของขั้นตอนวิธี SCRO ในการทดลองที่ 1 ทำให้การค้นคืนคำพิพาทศาสตร์ได้ประเด็นทางกฎหมายที่มีจำนวนน้อยกว่า เนื่องจากการคำพิพาทศาสตร์แต่ละฎีกาไม่สามารถมีคำค้นทั้งหมดที่ใช้ค้นคืนในครั้งเดียวได้ ทำให้ไม่พบฎีกาที่ต้องการ ดังนั้นการแบ่งกลุ่มคำค้นออกเป็น

จากภาพที่ 21 ตัวอย่างของการขยายคำค้น ของออนโทโลยีมรดกครั้งที่ 1 เช่น คำค้น “มรดก” กลุ่มคำที่ขยายได้คือ “มรดก” “ทายาท” “ตกทอด” “ผู้จัดการมรดก” “เจ้าหนี้” เป็นต้น



ภาพที่ 22 ออนโทโลยีมรดกครั้งที่ 2

จากภาพที่ 22 ออนโทโลยีมรดกครั้งที่ 2 “ทายาท” มีคำขยายคือ “ผู้สืบสิทธิ” “ผู้มีส่วนได้เสีย” “ผู้รับมรดก” เป็นต้น

3.3 ออนโทโลยีตัวเงิน

โครงสร้างของออนโทโลยีตัวเงินสำหรับขั้นตอนวิธี SCRO II สร้างจากคำพิพากษาศาลฎีกา และประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน มาตรา 898 ถึงมาตรา 1001 จำนวน 200 คำพิพากษาโครงสร้างของออนโทโลยีระดับที่ 1 ประกอบด้วยโนด 6 โนดและความสัมพันธ์ 29 ความสัมพันธ์ และระดับที่ 2 ประกอบด้วยโนด 2 โนดและความสัมพันธ์ 5 ความสัมพันธ์ โดยปรับโครงสร้างจากออนโทโลยีตัวเงินของขั้นตอนวิธี SCRO I แสดงดังภาพที่ 22

จากภาพที่ 24 ออนไลน์ด้วยตัวเงินระดับที่ 2 “ผู้ทรง” มีคำขยายคือ “ผู้ทรงเจ็ด” “ผู้ถือ” “ผู้รับเงิน” เป็นต้น

3.4 ขั้นตอนวิธี SCRO II

ผู้วิจัยปรับปรุงขั้นตอนของการท่องเที่ยวในออนไลน์ของขั้นตอนวิธี SCRO I เพื่อให้สามารถใช้ออนไลน์ที่มีการปรับโครงสร้างเป็น 2 ระดับได้ในการขยายคำค้น และรวมขั้นตอนในการเลือกคำพินิจ เพื่อลดความซับซ้อนของขั้นตอนวิธี SCRO I โดยตั้งชื่อว่าขั้นตอนวิธี SCRO II

ขั้นตอนวิธี SCRO II แสดงดังตารางที่ 4 โดยมีการปรับปรุงขั้นตอนขยายคำค้นโดยใช้ ออนไลน์ 2 ระดับ ขั้นตอนการดึงกลุ่มของคำค้น ขั้นตอนการค้นคืนเอกสาร ขั้นตอนการเรียงเอกสาร และขั้นตอนการตรวจสอบเอกสารที่ซ้ำกัน

จากตารางที่ 4 ข้อมูลเข้าคือคำค้นและออนไลน์ ในการทำงานขั้นตอนวิธี SCRO II มีการเก็บกลุ่มของคำค้นย่อย จากโครงสร้างของออนไลน์ จากนั้นนำคำค้นแต่ละกลุ่มไปค้นคืนคำพินิจทางสถิติ และเลือกเอกสารลำดับที่ 1 ของแต่ละกลุ่มคำค้นย่อย โดยตรวจสอบไม่ให้มีหมายเลขคำพินิจทางสถิติเดียวกัน ซึ่งมีขั้นตอนโดยละเอียดดังต่อไปนี้

ตารางที่ 4 ขั้นตอนวิธี SCRO II

Algorithm: SCRO II

Input: Ontology, keyword

Begin

Word = {}

Query = {}

ConceptType = {NODE, RELATION}

SearchConcept (Ontology, keyword)

Word = OntologyTraversal (keyword)

For Each w_i in Word Do

If ConceptType(keyword) = NODE and ConceptType(w_i) = NODE

If FoundExtendedOntology(w_i)

ExtendedWord = ExtendedOntologyTraversal()

If ConceptType(w_i) = RELATION and ConceptType(w_i) = RELATION

If FoundExtendedOntology(w_i)

ExtendedWord = ExtendedOntologyTraversal()

End For

Query = QueryGenerate(Word, ExtendedWord)

For each q_i in Query

D_i = DocumentRetrieval (q_i)

Ranking(D_i)

End For

CombineResults = checkDuplicate(D)

End

Output: CombineResults

1) คำค้นคือคำที่ใช้ในการค้นคืนคำพิพาทศาลฎีกา

2) ขั้นตอนการขยายคำค้นโดยใช้ออนโทโลยีคือขั้นตอนการขยายคำค้นจากโครงสร้างความสัมพันธ์ของออนโทโลยี โดยออนโทโลยีใช้ออนโทโลยี 2 ระดับ ออนโทโลยีระดับ 1 แสดงถึงความรู้โดยทั่วไปของกฎหมาย โดยโนดหรือความสัมพันธ์มี "*" ติดกับคำนั้น เพื่อแสดงถึงคำนั้นสามารถขยายไปยังออนโทโลยีระดับถัดไปได้ ขั้นตอนการขยายคำค้น มี 2 เงื่อนไขคือ

- ถ้ากรณีที่คำค้นพบในโนด การขยายคำจะขยายไปที่กลุ่มของคำที่มีความหมายเหมือนกัน

- ถ้ากรณีที่คำค้นพบในความสัมพันธ์ การขยายคำจะขยายไปที่กลุ่มของคำที่มีความสัมพันธ์กัน โดยจากภาพที่ 10 ตัวอย่างกลุ่มคำที่ได้จากการขยายคำค้นโดยใช้ออนโทโลยี 2 ระดับ เมื่อให้ r_1 เป็นคำค้น ได้โนดคือ $\{A, B\}$ กลุ่มความสัมพันธ์คือ $\{r_1^*, r_2, r_3, r_4\}$ และมี * อยู่ที่ r_1 ดังนั้นสามารถขยายคำ r_1 จากออนโทโลยีระดับที่ 2 ต่อไปได้คือ $\{word1, word2, word3\}$ ทำให้ได้กลุ่มของคำค้นให้คือ $\{A, B, r_1, word1, word2, word3, r_2, r_3, r_4\}$

3) กลุ่มของคำค้นคือกลุ่มของคำค้นที่ได้จากขั้นตอนการขยายคำค้น โดยโครงสร้างของกลุ่มของคำแต่ละกลุ่มประกอบด้วย โนดที่สัมพันธ์กับคำหลัก และความสัมพันธ์ที่เชื่อมระหว่างโนดและคำที่มีความสัมพันธ์กับคำหลัก

4) Legal Document คือการนำกลุ่มของคำค้นที่ได้จากการขยายคำค้น ค้นคืนคำพิพาทศาลฎีกา โดยใช้ค่าน้ำหนัก *TFIDF* จากสมการ (2)

5) Ranking & Checking Document คือการเรียงลำดับของคำพิพาทศาลฎีกาที่ค้นคืนออกมาได้ โดยทำการเรียงน้ำหนักของเอกสารจากมากไปน้อย และตรวจสอบคำพิพาทลำดับที่ 1 ของแต่ละกลุ่มไม่ให้ซ้ำกัน ซึ่งถ้าซ้ำกันจะให้กลุ่มที่มีน้ำหนักน้อย เลือกเอกสารลำดับถัดไปมาแทน

6) List of document คือการรวมผลลัพธ์ของทุกๆ คำพิพาทศาลฎีกาที่ได้ทำการตรวจสอบไม่ให้ซ้ำกันแล้ว แสดงผลลัพธ์ให้กับผู้ใช้

ผลและวิจารณ์

ผล

งานวิจัยนี้ได้แบ่งการทดลองออกเป็น 3 การทดลอง โดยแต่ละการทดลองผู้วิจัยได้ทำการปรับปรุงขั้นตอนวิธี และโครงสร้างของออนโทโลยี เพื่อเปรียบเทียบประสิทธิภาพ และผลลัพธ์ของแต่ละขั้นตอนวิธี

การทดลองที่ 1

การทดลองนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธีที่ผู้วิจัยได้เสนอคือขั้นตอนวิธี SCRO I และ SCRO II กับขั้นตอนวิธีที่ผู้วิจัยได้จำลองระบบสืบค้นในปัจจุบันคือขั้นตอนวิธี Baseline เพื่อแสดงประสิทธิภาพของการค้นคืนคำพิพากษาศาลฎีการะหว่างการขยายคำค้นจากออนโทโลยี 2 ชั้นและออนโทโลยี 1 ชั้นและเปรียบเทียบค่าความหลากหลายของระบบสืบค้นคำพิพากษาศาลฎีกาในปัจจุบันระหว่างขั้นตอนวิธี Deka I และ Deka II

1. ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองรวบรวมจากคำพิพากษาศาลฎีกาที่เกี่ยวข้องกับคดีมรดก มาตรา 1599-1754 จำนวน 248 คำพิพากษา และคดีที่เกี่ยวข้องกับตัวเงินตัวเงิน มาตรา 898-1001 จำนวน 493 คำพิพากษา

ออนโทโลยีที่ใช้คือออนโทโลยีมรดก และออนโทโลยีตัวเงิน ผู้เชี่ยวชาญสร้างจากคำพิพากษาศาลฎีกา และประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน มาตรา 898 ถึงมาตรา 1001 และออนโทโลยีมรดก สร้างจากคำพิพากษาศาลฎีกาและประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 6 มรดก มาตรา 1599 ถึงมาตรา 1754

2. ขั้นตอนวิธีที่ใช้ในการทดลอง

การทดลองที่ 1 ใช้ค่าความหลากหลายของประเด็นทางกฎหมายในการวัดประสิทธิภาพ โดยเปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธี SCRO I, SCRO II และ Baseline และเปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธี Deka I และ Deka II โดยแต่ละขั้นตอนวิธีมีรายละเอียดดังนี้

2.1 ขั้นตอนวิธี SCRO I ใช้กลุ่มคำค้นย่อยจากการขยายคำค้น โดยขยายความสัมพันธ์จาก ออนโทโลยี 1 ระดับ ในการค้นคำพิพากษาศาลฎีกาตามขั้นตอนวิธี SCRO I

2.2 ขั้นตอนวิธี SCRO I ใช้กลุ่มคำค้นย่อยจากการขยายคำค้น โดยขยายความสัมพันธ์จาก ออนโทโลยี 2 ระดับ ในการค้นคำพิพากษาศาลฎีกาตามขั้นตอนวิธี SCRO I

2.3 ขั้นตอนวิธี Baseline ใช้คำค้นเพียงคำเดียวในการค้นคำพิพากษาศาลฎีกาที่ผู้วิจัย ได้จำลองระบบไว้

2.4 ขั้นตอนวิธี Deka II ใช้กลุ่มคำค้นย่อยจากการขยายคำค้น โดยขยายความสัมพันธ์จาก ออนโทโลยี 2 ระดับ ในการค้นคำพิพากษาศาลฎีกาจากระบบสืบค้นคำพิพากษาในปัจจุบัน โดยแต่ละคำเชื่อมด้วยคำว่า “และ”

2.5 ขั้นตอนวิธี Deka I ใช้คำค้นเพียงคำเดียวในการค้นคำพิพากษาศาลฎีกาจากระบบ สืบค้นคำพิพากษาในปัจจุบัน

3. ผลการทดลอง

ผลการทดลองที่ 1 ความสัมพันธ์ในออนโทโลยีใช้แทนประเด็นทางกฎหมาย โดยโนดและความสัมพันธ์ของออนโทโลยีสามารถขยายคำไปสู่ออนโทโลยีระดับที่ 2 ได้ ซึ่งโนดแสดงถึงคำที่มีความหมายเหมือนกัน และความสัมพันธ์แสดงถึงความหลากหลายของประเด็นทางกฎหมาย โดย ออนโทโลยีที่ใช้ในการทดลองคือออนโทโลยีมรดก และออนโทโลยีตัวเงิน ข้อมูลที่ใช้คือคำพิพากษาศาลฎีกามรดกและตัวเงิน

ผลการทดลองที่ 1 บนชุดข้อมูลมรดกเปรียบเทียบจำนวนประเด็นความหลากหลายของ
ขั้นตอนวิธี SCRO II, SCRO I และ Baseline โดยใช้คำในการทดลอง 12 คำ แสดงดังตารางที่ 5

ตารางที่ 5 ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของประเด็นทางกฎหมายของ
ขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลมรดก

คำค้น	จำนวนประเด็น ทั้งหมด	จำนวนประเด็นความหลากหลายแต่ละขั้นตอน		
		Baseline	SCRO I	SCRO II
มรดก	19	8	8	14
ผู้จัดการมรดก	19	6	10	10
ทายาท	17	9	9	14
ผู้ตาย	4	1	3	4
พยาน	3	2	3	3
พินัยกรรม	9	2	9	6
เจ้าหนี้	7	1	3	4
กำจัดมิให้รับมรดก	6	2	2	4
สละมรดก	3	1	2	2
พินัยกรรมไม่สมบูรณ์	4	2	2	3
ชำระหนี้	3	1	1	2
ตั้งผู้จัดการมรดก	5	2	3	3

จากตารางที่ 5 จำนวนประเด็นทั้งหมดได้จากความสัมพันธ์ของออนโทโลยีมรดก โดย 1
ประเด็นคือ 1 ความสัมพันธ์ระหว่าง โนด 2 โนดและเมื่อใช้แต่ละขั้นตอนค้นคืนแล้วผู้เชี่ยวชาญเป็นผู้
พิจารณาความหลากหลายของแต่ละเอกสาร

ขั้นตอนวิธี SCRO II สามารถค้นคืนจำนวนประเด็นทางกฎหมายได้มากกว่าขั้นตอนวิธี
SCRO I ในส่วนของคำว่า “มรดก” “ทายาท” “ผู้ตาย” “เจ้าหนี้” “กำจัดมิให้รับมรดก” “พินัยกรรม
ไม่สมบูรณ์” “ชำระหนี้” และ “ตั้งผู้จัดการมรดก” หากพิจารณาความแตกต่างระหว่างคำ 2 คำคือ
“มรดก” และ “ทายาท” ซึ่ง 2 คำนี้เป็นคำที่มีความหมายกว้าง ไม่เฉพาะเจาะจง ดังนั้นการค้นคืน
ประเด็นความหลากหลายจากคำทั้ง 2 จึงเป็นเรื่องที่ยาก พิจารณาคำว่า “มรดก” ขั้นตอนวิธี SCRO II

สามารถค้นคืน ได้ฎีกาที่มีประเด็นความหลากหลาย 14 ฎีกา ขณะที่ขั้นตอนวิธี SCRO I สามารถค้นคืนได้ฎีกาที่มีประเด็นความหลากหลาย 8 ฎีกา ส่วนคำว่า “ทายาท” ขั้นตอนวิธี SCRO II สามารถค้นคืนได้ฎีกาที่มีประเด็นความหลากหลาย 14 ฎีกา และขั้นตอนวิธี SCRO I สามารถค้นคืนได้ฎีกาที่มีประเด็นความหลากหลาย 9 ฎีกา

ในการวัดประสิทธิภาพของการทดลองที่ 1 ใช้ค่าความหลากหลายของประเด็นทางกฎหมายจากสมการที่ (7) โดยเปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธี SCRO II, SCRO I และ Baseline ดังตารางที่ 6

ตารางที่ 6 ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลมรดก

คำค้น	ค่าความหลากหลายของประเด็นทางกฎหมาย		
	Baseline	SCRO I	SCRO II
มรดก	0.42	0.42	0.74
ผู้จัดการมรดก	0.32	0.53	0.53
ทายาท	0.53	0.53	0.82
ผู้ตาย	0.25	0.75	1.00
พยาน	0.67	1.00	1.00
พินัยกรรม	0.22	1.00	0.67
เจ้าหนี้	0.14	0.43	0.57
กำจัดมิให้รับมรดก	0.33	0.33	0.67
สละมรดก	0.33	0.67	0.67
พินัยกรรมไม่สมบูรณ์	0.50	0.50	0.75
ชำระหนี้	0.33	0.33	0.67
ตั้งผู้จัดการมรดก	0.40	0.60	0.60
ค่าเฉลี่ย	0.37	0.59	0.72

จากตารางที่ 6 พบว่าขั้นตอนวิธี SCRO II ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายสูงสุดเท่ากับ 72% ขั้นตอนวิธี SCRO I ให้ค่าความหลากหลายของประเด็นทางกฎหมาย

เท่ากับ 59% และขั้นตอนวิธี Baseline ให้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 37% โดยขั้นตอนวิธี SCRO II สามารถเพิ่มค่าความหลากหลายของประเด็นทางกฎหมายได้มากขึ้นใน ส่วนของคำค้นที่เป็นความสัมพันธ์คือคำว่า “ก่อกวนให้รับมรดก” “พินัยกรรมไม่สมบูรณ์” และ “ชำระหนี้” เนื่องจากการขยายคำค้นออกเป็นกลุ่มย่อย และมีการขยายคำในส่วนของโน้ตและความสัมพันธ์ระดับชั้นที่ 2

ในผลการทดลองที่ 1 บนชุดข้อมูลมรดกผู้วิจัยได้เปรียบเทียบจำนวนประเด็นความ หลากหลายของขั้นตอนวิธี Deka I และ Deka II ซึ่งเป็นการค้นคืนคำพิพากษาศาลฎีกาในระบบจริง โดยใช้คำในการทดลอง 12 คำ ซึ่งจำนวนประเด็นความหลากหลายของแต่ละขั้นตอนวิธี แสดงดัง ตารางที่ 7

ตารางที่ 7 ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของขั้นตอนวิธี Deka I และ Deka II บนชุดข้อมูลมรดก

คำค้น	จำนวนประเด็น ทั้งหมด	จำนวนประเด็นความหลากหลาย	
		Deka I	Deka II
มรดก	19	7	10
ผู้จัดการมรดก	19	4	6
ทายาท	17	3	8
ผู้ตาย	4	1	2
พยาน	3	1	1
พินัยกรรม	9	3	2
เจ้าหนี้	7	3	2
ก่อกวนให้รับมรดก	6	2	6
สละมรดก	3	1	3
พินัยกรรมไม่สมบูรณ์	4	2	1
ชำระหนี้	3	2	1
ตั้งผู้จัดการมรดก	5	1	1

จากตารางที่ 7 จำนวนประเด็นทั้งหมดได้จากความสัมพันธ์ของออนโทโลยีมรดก โดย 1 ประเด็นคือ 1 ความสัมพันธ์ระหว่างโนด 2 โนดและเมื่อใช้แต่ละขั้นตอนกันขึ้นแล้วผู้เชี่ยวชาญเป็นผู้พิจารณาความหลากหลายของแต่ละเอกสาร แสดงให้เห็นว่าขั้นตอนวิธี Deka II สามารถค้นคืนจำนวนประเด็นทางกฎหมายได้มากกว่าขั้นตอนวิธี Deka I ในส่วนของคำว่า “มรดก” “ผู้จัดการมรดก” “ทายาท” “ผู้ตาย” “กํารั้ดมิให้รับมรดก” และ “สละมรดก” โดยคำเหล่านี้เป็นทั้งคำในโนดและความสัมพันธ์ แสดงให้เห็นว่าการขยายคำค้นจากออนโทโลยีสามารถเพิ่มประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายให้กับระบบสืบค้นในปัจจุบันได้

ในการวัดประสิทธิภาพของการทดลองที่ 3 ใช้ค่าความหลากหลายของประเด็นทางกฎหมายจากสมการที่ (7) ซึ่งจากตารางที่ 11 สามารถเปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Deka I และ Deka II ได้ดังตารางที่ 8

ตารางที่ 8 ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Deka II และ Deka I บนชุดข้อมูลมรดก

คำค้น	ค่าความหลากหลายของประเด็นทางกฎหมาย	
	Deka I	Deka II
มรดก	0.37	0.53
ผู้จัดการมรดก	0.21	0.32
ทายาท	0.18	0.47
ผู้ตาย	0.25	0.50
พยาน	0.33	0.33
พินัยกรรม	0.33	0.22
เจ้าหนี้	0.43	0.29
กํารั้ดมิให้รับมรดก	0.33	1.00
สละมรดก	0.33	1.00
พินัยกรรมไม่สมบูรณ์	0.50	0.25
ชำระหนี้	0.67	0.33
ตั้งผู้จัดการมรดก	0.20	0.20
ค่าเฉลี่ย	0.34	0.45

จากตารางที่ 8 การขยายคำค้นแล้วนำไปค้นคืนในขั้นตอนวิธี Deka II สามารถค้นคืนคำว่า “กำจัดมิให้รับมรดก” และ “สละมรดก” ได้ความหลากหลายสูงสุดเท่ากับ 100% เนื่องจากเป็นกลุ่มคำที่ขยายจากออนโทโลยีและนำไปค้นคืนในระบบจริง ทำให้ผลการค้นคืนมีค่าความหลากหลายของประเด็นทางกฎหมายสูงกว่าขั้นตอนวิธี Deka I ที่ใช้คำค้นในการค้นคืนโดยตรง

ค่าความหลากหลายของประเด็นทางกฎหมายเฉลี่ยของ Deka II เท่ากับ 45% และ Deka I เท่ากับ 34% แสดงให้เห็นว่าการขยายคำค้นโดยใช้ออนโทโลยี สามารถเพิ่มประสิทธิภาพของการค้นคืนคำพิพากษาศาลฎีกาให้มีค่าความหลากหลายของประเด็นทางกฎหมายที่เพิ่ม

ในการพิจารณาค่าความถี่ของคำ ลักษณะของคำ และค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO II, SCRO I และ Baseline แสดงดังตารางที่ 9

ตารางที่ 9 การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลมรดก

คำ	ความถี่ ของคำ	ลักษณะ ของคำ	ค่าความหลากหลายของประเด็นทางกฎหมาย		
			Baseline	SCRO I	SCRO II
มรดก	1201	นาม	0.42	0.42	0.74
ผู้ตาย	516	นาม	0.25	0.75	1.00
ผู้จัดการมรดก	364	นาม	0.32	0.53	0.53
พินัยกรรม	347	นาม	0.22	1.00	0.67
ทายาท	340	นาม	0.53	0.53	0.82
เจ้าหนี้	69	นาม	0.14	0.43	0.57
พยาน	62	นาม	0.67	1.00	1.00
ชำระหนี้	62	กริยา	0.33	0.33	0.67
สละมรดก	16	กริยา	0.33	0.67	0.67
ตั้งผู้จัดการมรดก	13	กริยา	0.40	0.60	0.60
พินัยกรรมไม่สมบูรณ์	8	กริยา	0.50	0.50	0.75
กำจัดมิให้รับมรดก	6	กริยา	0.33	0.33	0.67

จากตารางที่ 9 ค่าความถี่ของคำเรียงจากมากไปน้อย และแบ่งกลุ่มตามลักษณะของคำ โดยในการทดลองที่ 1 บนชุดข้อมูลมรดกประเด็นต่างๆ ของกฎหมายทั้งหมดได้จากความสัมพันธ์ทั้งหมดที่เชื่อมระหว่างโนดที่มีความสัมพันธ์กัน ซึ่งคำที่มีค่าความถี่สูง ส่งผลให้ประเด็นต่างๆ ของกฎหมายมีค่าสูง จากผลการทดลอง ค่าความถี่ของคำสูงสุดคือคำว่า “มรดก” ซึ่งขั้นตอนวิธี SCRO II คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.74 ขั้นตอนวิธี SCRO I คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.42 และขั้นตอนวิธี Baseline คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.42

ลักษณะของคำไม่มีผลต่อการค้นหาของขั้นตอนวิธี SCRO II และ SCRO I แต่มีผลต่อประเด็นทั้งหมดของกฎหมาย เนื่องจากน้ำหนักที่ใช้ในการคำนวณคือค่า *TFIDF* ซึ่ง *IDF* เป็นส่วนในการกระจายค่าน้ำหนักของคำ ไม่ให้ค่าความถี่ของคำที่สูงมีผลต่อการค้นคืนเอกสาร

ตารางที่ 10 การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความหลากหลายของประเด็นทางกฎหมาย ของขั้นตอนวิธี Deak I และ Deak II บนชุดข้อมูลมรดก

คำ	ความถี่ ของคำ	ลักษณะ ของคำ	ค่าความหลากหลายของประเด็นทางกฎหมาย	
			Deka I	Deka II
มรดก	1201	นาม	0.37	0.53
ผู้ตาย	516	นาม	0.25	0.50
ผู้จัดการมรดก	364	นาม	0.21	0.32
พินัยกรรม	347	นาม	0.33	0.22
ทายาท	340	นาม	0.18	0.47
เจ้าหนี้	69	นาม	0.43	0.29
พยาน	62	นาม	0.33	0.33
ชำระหนี้	62	กริยา	0.67	0.33
สละมรดก	16	กริยา	0.33	1.00
ตั้งผู้จัดการมรดก	13	กริยา	0.20	0.20
พินัยกรรมไม่สมบูรณ์	8	กริยา	0.50	0.25
กำจัดมิให้รับมรดก	6	กริยา	0.33	1.00

จากตารางที่ 10 ค่าความถี่ของคำเรียงจากมากไปน้อย และแบ่งกลุ่มตามลักษณะของคำ โดยในการทดลองที่ 3 บนชุดข้อมูลมรดกประเด็นต่างๆ ของกฎหมายทั้งหมดได้จากความสัมพันธ์ทั้งหมดที่เชื่อมระหว่างโนดที่มีความสัมพันธ์กัน ซึ่งคำที่มีค่าความถี่สูง ส่งผลให้ประเด็นต่างๆ ของกฎหมายมีค่าสูง จากผลการทดลอง ค่าความถี่ของคำสูงสุดคือคำว่า “มรดก” ซึ่งขั้นตอนวิธี Deka II คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.53 และขั้นตอนวิธี Deka I คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.37 ส่วนคำที่มีความถี่น้อยที่สุดคือ คำว่า “กำจัดมิให้รับมรดก” ขั้นตอนวิธี Deka II คำนวณได้ค่าความหลากหลาย 1.00 และขั้นตอนวิธี Deka I คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.33

ผลการทดลองที่ 1 บนชุดข้อมูลตัวเงินเปรียบเทียบจำนวนประเด็นความหลากหลายของขั้นตอนวิธี SCRO II, SCRO I และ Baseline โดยใช้ค่าในการทดลอง 12 คำ แสดงดังตารางที่ 11

ตารางที่ 11 ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของขั้นตอนวิธี SCRO II, SCRO I และ Baseline บนชุดข้อมูลตัวเงิน

คำค้น	จำนวนประเด็นทั้งหมด	จำนวนความประเด็นความหลากหลายแต่ละขั้นตอน		
		Baseline	SCRO I	SCRO II
เช็ค	13	2	10	9
ตัวแลกเงิน	5	1	3	4
ตัวสัญญาใช้เงิน	9	3	7	6
ผู้ทรง	12	3	4	8
ธนาคาร	5	2	5	4
หนี้	13	6	10	9
สลักหลัง	3	2	2	3
ดอกเบี้ย	2	1	2	2
โอน	3	1	1	1
ไล่เบี้ย	2	1	1	2
อาวัล	6	1	4	5
สั่งจ่าย	2	1	1	2

จากตารางที่ 11 จำนวนประเด็นทั้งหมดได้จากความสัมพันธ์ของออนโทโลยีตัวเงิน โดย 1 ประเด็นคือ 1 ความสัมพันธ์ระหว่าง โหนด 2 โหนดและเมื่อใช้แต่ละขั้นตอนกันแล้วผู้เชี่ยวชาญเป็นผู้พิจารณาความหลากหลายของแต่ละเอกสาร

ขั้นตอนวิธี SCRO II สามารถค้นคืนจำนวนประเด็นทางกฎหมายได้เพิ่มขึ้น เมื่อเปรียบเทียบกับขั้นตอนวิธี SCRO I ในส่วนของคำว่า “ผู้ทรง” “สลักหลัง” “ไล่เบี้ย” “อาวัล” และ “สั่งจ่าย” ซึ่งเป็นคำส่วนใหญ่เป็นคำกริยาที่ใช้แทนประเด็นต่างๆ ของกฎหมาย

ในการวัดประสิทธิภาพของการทดลองที่ 1 ใช้ค่าความหลากหลายของประเด็นทางกฎหมายจากสมการที่ (7) ซึ่งจากตารางที่ 15 สามารถเปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II ดังตารางที่ 12

ตารางที่ 12 ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลตัวเงิน

คำค้น	ค่าความหลากหลายของประเด็นทางกฎหมาย		
	Baseline	SCRO I	SCRO II
เช็ค	0.15	0.77	0.69
ตั๋วแลกเงิน	0.20	0.60	0.80
ตั๋วสัญญาใช้เงิน	0.33	0.78	0.67
ผู้ทรง	0.25	0.33	0.67
ธนาคาร	0.40	1.00	0.80
หนี้	0.46	0.77	0.69
สลักหลัง	0.67	0.67	1.00
ดอกเบี้ย	0.50	1.00	1.00
โอน	0.33	0.33	0.33
ไล่เบี้ย	0.50	0.50	1.00
อาวัล	0.17	0.67	0.83
สั่งจ่าย	0.50	0.50	1.00
ค่าเฉลี่ย	0.37	0.66	0.79

เมื่อเปรียบเทียบระหว่างขั้นตอนวิธี SCRO II และ SCRO I คำว่า “สลักหลัง” “ดอกเบี้ย” “ไล่เบี้ย” และ “สั่งจ่าย” ของขั้นตอนวิธี SCRO II คั่นกันได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 100% ส่วนคำว่า “ธนาคาร” และ “ดอกเบี้ย” ของขั้นตอนวิธี SCRO I คั่นกันได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 100% เช่นกัน ซึ่งถ้าพิจารณาเฉพาะค่าความหลากหลายของประเด็นทางกฎหมายในส่วนของความสัมพันธ์ขั้นตอนวิธี SCRO II มีประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายที่ดีกว่าขั้นตอนวิธี SCRO I คือคำว่า “สลักหลัง” “อาวัล” “สั่งจ่าย”

ในผลการทดลองที่ 3 บนชุดข้อมูลมรดกผู้วิจัยได้เปรียบเทียบจำนวนประเด็นความหลากหลายของขั้นตอนวิธี Deka I และ Deka II ซึ่งเป็นการคั่นคำพิพากษาศาลฎีกาในระบบจริง โดยใช้คำในการทดลอง 12 คำ ซึ่งจำนวนประเด็นความหลากหลายของแต่ละขั้นตอนวิธี แสดงดังตารางที่ 13

ตารางที่ 13 ผลการทดลองที่ 1 แสดงจำนวนประเด็นที่หลากหลายของขั้นตอนวิธี Deka I และ Deka II บนชุดข้อมูลตัวเงิน

คำค้น	จำนวนประเด็นทั้งหมด	จำนวนความประเด็นความหลากหลาย	
		Deka I	Deka II
เช็ค	13	6	7
ตั๋วแลกเงิน	5	1	3
ตั๋วสัญญาใช้เงิน	9	4	7
ผู้ทรง	12	3	4
ธนาคาร	5	3	4
หนี้	13	8	5
สลักหลัง	3	1	2
ดอกเบี้ย	2	0	1
โอน	3	0	1
ไล่เบี้ย	2	0	1
อาวัล	6	1	3
สั่งจ่าย	2	0	1

จากตารางที่ 13 จำนวนประเด็นทั้งหมดได้จากความสัมพันธ์ของออนโทโลยีมรดก โดย 1 ประเด็นคือ 1 ความสัมพันธ์ระหว่างโนด 2 โนดและเมื่อใช้แต่ละขั้นตอนกันคืนแล้วผู้เชี่ยวชาญเป็นผู้พิจารณาความหลากหลายของแต่ละเอกสาร พบว่าจำนวนประเด็นความหลากหลายทั้งหมดขั้นตอนวิธี Deka II สามารถค้นคืนจำนวนประเด็นทางกฎหมายได้เพิ่มขึ้นยกเว้นคำว่า “หนี้” เมื่อเปรียบเทียบกับขั้นตอนวิธี Deka I เนื่องจากขั้นตอนวิธี Deka II ใช้กลุ่มของคำค้นที่ขยายจากออนโทโลยี ก่อนนำไปค้นคืนคำพิพากษาศาลฎีกา ซึ่งบอกถึงประเด็นต่างๆ ทางกฎหมายที่มีความหลากหลาย

ในการวัดประสิทธิภาพของการทดลองที่ 1 ใช้ค่าความหลากหลายของประเด็นทางกฎหมายจากสมการที่ (7) ซึ่งจากตารางที่ 10 สามารถเปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Deka I และ Deka II ดังตารางที่ 14

ตารางที่ 14 ผลการทดลองที่ 1 แสดงค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Deka I และ Deka II บนชุดข้อมูลตัวเงิน

คำค้น	ค่าความหลากหลายของประเด็นทางกฎหมาย	
	Deka I	Deka II
เช็ก	0.46	0.54
ตัวแลกเงิน	0.20	0.60
ตัวสัญญาใช้เงิน	0.44	0.78
ผู้ทรง	0.25	0.33
ธนาคาร	0.60	0.80
หนี้	0.62	0.38
สลักหลัง	0.33	0.67
ดอกเบี้ย	0.00	0.50
โอน	0.00	0.33
ไล่เบี้ย	0.00	0.50
อาวัล	0.17	0.50
ตั้งจ่าย	0.00	0.50
ค่าเฉลี่ย	0.25	0.26

จากตารางที่ 14 การขยายคำค้นแล้วนำไปค้นคืนในขั้นตอนวิธี Deka II สามารถค้นคืนคำว่า “ธนาคาร” ได้ความหลากหลายสูงสุดเท่ากับ 80% เนื่องจากเป็นกลุ่มคำที่ขยายจากออนโทโลยีและนำไปค้นคืนในระบบจริง ทำให้ผลการค้นคืนมีค่าความหลากหลายของประเด็นทางกฎหมายสูงกว่าขั้นตอนวิธี Deka I ที่ให้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 60%

ขั้นตอนวิธี Deka I มีส่วนของค่าความหลากหลายของประเด็นทางกฎหมายเป็นศูนย์กลางคือคำว่า “ดอกเบี้ย”, “โอน”, “ไต่เบี่ย” และ “สั่งจ่าย” เนื่องจากคำค้นที่ใช้เมื่อทำการค้นคืนแล้ว คำนั้นไม่ปรากฏในส่วนของย่อสั้น ซึ่งเป็นส่วนที่ใช้ในการพิจารณาของแต่ละประเด็น แต่คำนั้นปรากฏอยู่ในส่วนของย่อยาว เมื่อพิจารณาที่ความหมายโดยรวมแล้วคำพิพาทภาษานั้นก็ไม่มีประเด็นที่ตรงกับคำค้นนั้นเลย เช่น คำค้นนั้นเป็นคำที่ใช้ในการอ้างตัวอย่างของกฎหมาย ทำให้ผลการค้นคืนไม่ได้ประเด็นตามที่ต้องการ

ตารางที่ 15 การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II บนชุดข้อมูลตัวเงิน

คำ	ความถี่ ของคำ	ลักษณะ ของคำ	ค่าความหลากหลายของประเด็นทางกฎหมาย		
			Baseline	SCRO I	SCRO II
หนี้	2573	นาม	0.46	0.77	0.69
เช็ค	1720	นาม	0.15	0.77	0.69
ดอกเบี้ย	632	นาม	0.50	1.00	1.00
ธนาคาร	602	นาม	0.40	1.00	0.80
ตัวสัญญาใช้เงิน	122	นาม	0.33	0.78	0.67
ผู้ทรง	104	นาม	0.25	0.33	0.67
ตัวแลกเงิน	55	นาม	0.20	0.60	0.80
สั่งจ่าย	335	กริยา	0.50	0.50	1.00
โอน	265	กริยา	0.33	0.33	0.33
สลักหลัง	49	กริยา	0.67	0.67	1.00
ไต่เบี่ย	47	กริยา	0.50	0.50	1.00
อาวัล	35	กริยา	0.17	0.67	0.83

จากตารางที่ 15 ค่าความถี่ของคำเรียงจากมากไปน้อย และแบ่งกลุ่มตามลักษณะของคำ โดยในการทดลองที่ 2 บนชุดข้อมูลตัวเงินประเด็นต่างๆ ของกฎหมายทั้งหมดได้จากความสัมพันธ์ทั้งหมดที่เชื่อมระหว่างโนดที่มีความสัมพันธ์กัน ซึ่งคำที่มีค่าความถี่สูง ส่งผลให้ประเด็นต่างๆ ของกฎหมายมีค่าสูง จากผลการทดลอง ค่าความถี่ของคำสูงสุดคือคำว่า “หนี้” ซึ่งขั้นตอนวิธี SCRO I คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.77 และขั้นตอนวิธี Baseline คำนวณได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.38

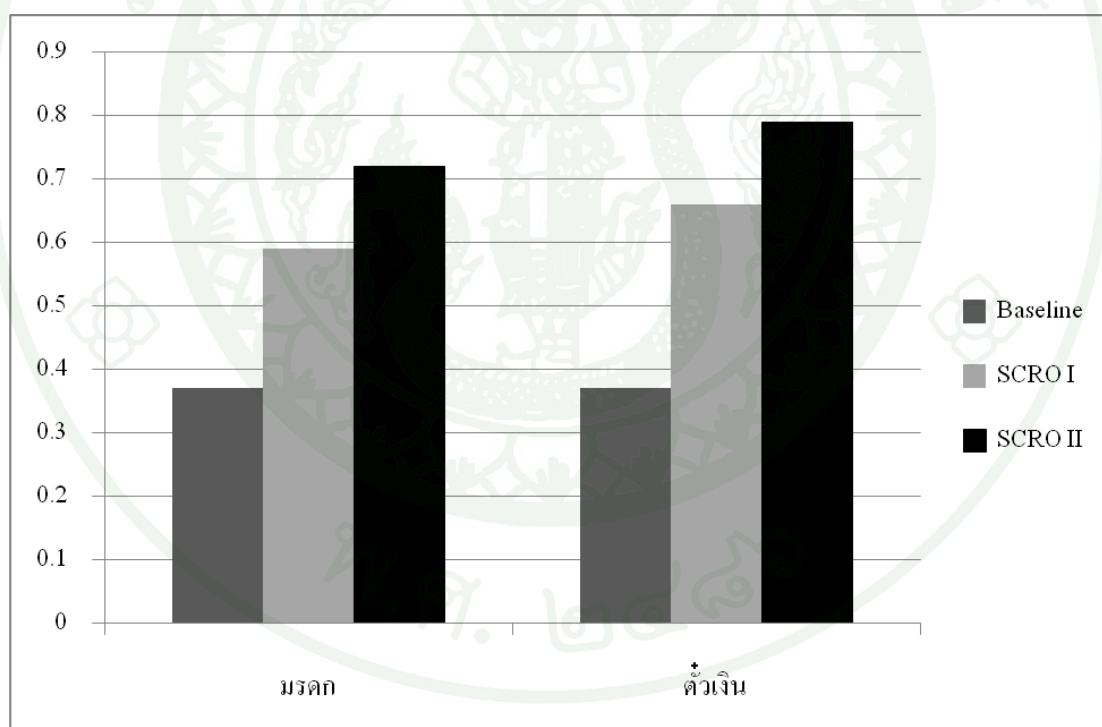
ลักษณะของคำไม่มีผลต่อการค้นหาของขั้นตอนวิธี SCRO II และ SCRO I แต่มีผลต่อประเด็นทั้งหมดของกฎหมาย เนื่องจากน้ำหนักที่ใช้ในการคำนวณคือค่า $TFIDF$ ซึ่ง IDF เป็นส่วนในการกระจายค่าน้ำหนักของคำ ไม่ให้ค่าความถี่ของคำที่สูงมีผลต่อการค้นคืนเอกสาร

ตารางที่ 16 การทดลองที่ 1 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ ค่าความหลากหลายของประเด็นทางกฎหมาย ของขั้นตอนวิธี Deka I และ Deka II บนชุดข้อมูลตัวเงิน

คำ	ความถี่ ของคำ	ลักษณะ ของคำ	ค่าความหลากหลายของประเด็นทางกฎหมาย	
			Deka I	Deka II
หนี้	2573	นาม	0.62	0.38
เช็ค	1720	นาม	0.46	0.54
ดอกเบี้ย	632	นาม	0.00	0.50
ธนาคาร	602	นาม	0.60	0.80
ตัวสัญญาใช้เงิน	122	นาม	0.44	0.78
ผู้ทรง	104	นาม	0.25	0.33
ตัวแลกเงิน	55	นาม	0.20	0.60
ส่งจ่าย	335	กริยา	0.00	0.50
โอน	265	กริยา	0.00	0.33
สลักหลัง	49	กริยา	0.33	0.67
ไล่เบี้ย	47	กริยา	0.00	0.50
อาวัล	35	กริยา	0.17	0.50

จากตารางที่ 16 ค่าความถี่ของคำเรียงจากมากไปน้อย และแบ่งกลุ่มตามลักษณะของคำ โดยในการทดลองที่ 1 คดีตัวเงินประเด็นต่างๆ ของกฎหมายทั้งหมดได้จากความสัมพันธ์ทั้งหมดที่เชื่อมระหว่างโนดที่มีความสัมพันธ์กัน ซึ่งคำที่มีค่าความถี่สูง ส่งผลให้ประเด็นต่างๆ ของกฎหมายมีค่าสูง จากผลการทดลอง ค่าความถี่ของคำสูงสุดคือคำว่า “หนี้” ซึ่งขั้นตอนวิธี Deka II ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.38 และขั้นตอนวิธี Deka I ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.62 ส่วนคำที่มีความถี่น้อยที่สุดคือ คำว่า “อาวัล” ขั้นตอนวิธี Deka II ค้นคืนได้ค่าความหลากหลายเท่ากับ 0.50 และขั้นตอนวิธี Deka I ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.17

จากผลการทดลองที่ 1 สามารถนำขั้นตอนวิธี Baseline, SCRO I และ SCRO II สร้างกราฟแสดงประสิทธิภาพของการค้นคืนค่าความหลากหลายของประเด็นทางกฎหมายได้ดังภาพที่ 25

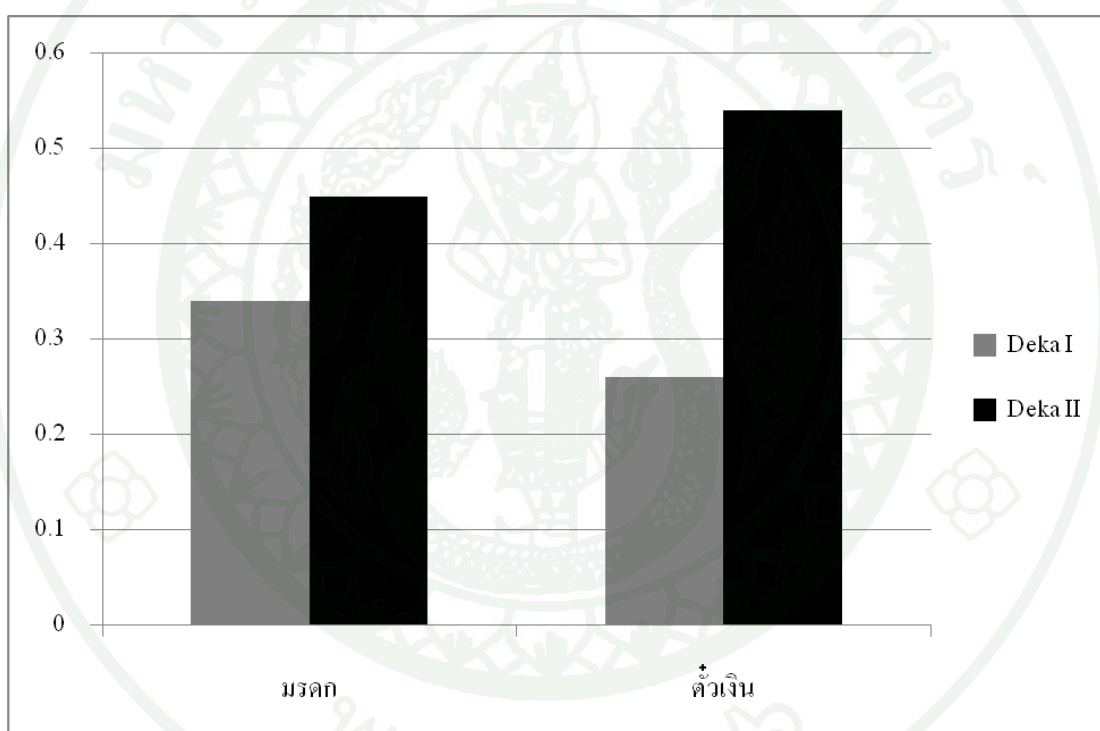


ภาพที่ 25 ค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline, SCRO I และ SCRO II

จากภาพที่ 25 ค่าความหลากหลายของประเด็นทางกฎหมายของแต่ละขั้นตอนวิธี โดยแนวตั้งคือค่าความหลากหลายของประเด็นทางกฎหมาย และแนวนอนคือขั้นตอนวิธีต่างๆ ผลการ

ทดลองแสดงให้เห็นว่าขั้นตอนวิธี SCRO II สามารถค้นคืนให้ค่าความหลากหลายของประเด็นทางกฎหมายสูงสุดทั้ง 2 ชุดข้อมูลทดสอบ รองลงมาคือขั้นตอนวิธี SCRO I แสดงให้เห็นว่าการแบ่งกลุ่มคำค้นออกเป็นกลุ่มย่อยๆ โดยใช้ความสัมพันธ์ของออนโทโลยี สามารถเพิ่มประสิทธิภาพของการค้นคืนคำพิพากษาศาลฎีกาให้มีความหลากหลายได้ มากกว่าการใช้คำค้นจากผู้ใช้เพียงคำเดียวในการค้นคืนเอกสาร

จากผลการทดลองที่ 1 สามารถนำขั้นตอนวิธี Deka I และ Deka II เขียนเป็นกราฟแสดงประสิทธิภาพของการค้นคืนค่าความหลากหลายของประเด็นทางกฎหมายได้ดังภาพที่ 26



ภาพที่ 26 ค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Deka I และ Deka II

จากภาพที่ 26 เมื่อนำคำค้นที่ได้จากการขยายคำค้นของขั้นตอนวิธี SCRO II และนำกลุ่มคำที่ได้ไปค้นคืนในระบบสืบค้นคำพิพากษาและคำสั่งคำร้องศาลฎีกา เชื่อมด้วยคำว่า “และ” เปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายกับการใช้คำค้นโดยตรงจากผู้ค้นคืนในระบบสืบค้นคำพิพากษาและคำสั่งคำร้องศาลฎีกา แสดงให้เห็นว่าขั้นตอนวิธี Deka II สามารถให้ค่าความหลากหลายของประเด็นทางกฎหมายสูงกว่าขั้นตอนวิธี Deka I ทั้งการทดลองมรดกและตัวเงิน เนื่องจากกลุ่มของคำค้นย่อยที่ได้จากการขยายคำค้นโดยใช้ออนโทโลยี แสดงถึงประเด็นต่างๆ

ทางกฎหมาย ทำให้ประสิทธิภาพของการค้นคืนในระบบจริงเพิ่มค่าความหลากหลายของประเด็นทางกฎหมายให้กับผลลัพธ์ในแต่ละลำดับ

การขยายคำค้นโดยใช้ออนโทโลยีของการทดลองที่ 1 ได้กลุ่มของคำค้นที่แต่ละกลุ่มแสดงถึงประเด็นต่างๆ ของกฎหมาย และมีการเพิ่มคำที่มีความสัมพันธ์กับความสัมพันธ์ที่สามารถขยายออกไปได้ในออนโทโลยีระดับที่ 2 ทำให้ได้กลุ่มของคำค้นที่มีประเด็นทางกฎหมายเพิ่มมากขึ้น ผลการทดลองแสดงให้เห็นว่าขั้นตอนวิธี SCRO II มีประสิทธิภาพมากกว่าขั้นตอนวิธี SCRO I และเมื่อทำการทดลองกับระบบจริง ผลการทดลองแสดงให้เห็นว่าขั้นตอนวิธี Deka II มีประสิทธิภาพที่ดีกว่าขั้นตอนวิธี Deka I

จากผลการทดลองของการทดลองที่ 1 ขั้นตอนวิธี SCRO II สามารถเพิ่มประสิทธิภาพของค่าความหลากหลายทางประเด็นทางกฎหมายในส่วน of คำที่เป็นความสัมพันธ์ได้ โดยการขยายคำไปยังออนโทโลยีระดับที่ 2 ทำให้ผลลัพธ์ของคำพิพากษาศาลฎีกามีประเด็นทางกฎหมายเพิ่มขึ้น ขณะที่ขั้นตอนวิธี SCRO I มีเพียงออนโทโลยีระดับเดียวซึ่งสามารถขยายกลุ่มของคำค้นที่เป็นโครงสร้างความสัมพันธ์หลักได้เท่านั้น ทำให้ประสิทธิภาพของการค้นคืนต่ำกว่าขั้นตอนวิธี SCRO II

ขั้นตอนวิธี Deka II สามารถค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายได้มากกว่าขั้นตอนวิธี Deka I เนื่องจาก ขั้นตอนวิธี Deka II ใช้กลุ่มของคำค้นที่ได้จากการขยายคำค้นโดยออนโทโลยีออกเป็นกลุ่มย่อยๆ โดยแต่ละกลุ่มแทนประเด็นทางกฎหมายที่มีความหลากหลาย ทำให้ประสิทธิภาพของการค้นคืนคำพิพากษาศาลฎีกาของขั้นตอนวิธี Deka II เพิ่มขึ้น

การทดลองที่ 2

วัตถุประสงค์ของการทดลองนี้เพื่อเปรียบเทียบประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธีที่ผู้วิจัยได้เสนอคือขั้นตอนวิธี SCRO I กับขั้นตอนวิธีที่ผู้วิจัยได้จำลองระบบสืบค้นในปัจจุบันคือขั้นตอนวิธี Baseline เพื่อแสดงประสิทธิภาพของการค้นคืนคำพิพากษาศาลฎีกาจากออนโทโลยี 1 ชั้น

1. ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองรวบรวมจากคำพิพากษาศาลฎีกาที่เกี่ยวข้องกับคดีมรดก มาตรา 1599-1754 จำนวน 248 คำพิพากษา และคดีที่เกี่ยวข้องกับตัวเงินตัวเงิน มาตรา 898-1001 จำนวน 493 คำพิพากษา

อนโทโลยีที่ใช้คืออนโทโลยีมรดก และอนโทโลยีตัวเงินจากผู้เชี่ยวชาญสร้างจากคำพิพากษาศาลฎีกา และประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน มาตรา 898 ถึงมาตรา 1001 และอนโทโลยีมรดก สร้างจากคำพิพากษาศาลฎีกาและประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 6 มรดก มาตรา 1599 ถึงมาตรา 1754

2. ขั้นตอนวิธีที่ใช้ในการทดลอง

ในการทดลอง ทดลองกับคำพิพากษาศาลฎีกาที่เกี่ยวข้องกับคดีมรดก และ ตัวเงิน โดยคดีมรดกใช้คำค้นทั้งหมด 10 คำ และ คดีตัวเงินใช้คำค้นทั้งหมด 12 คำ ซึ่งเป็นคำที่ปรากฏในโนดและความสัมพันธ์ของอนโทโลยี ผลการทดลองเปรียบเทียบกับขั้นตอนดังต่อไปนี้

2.1 ขั้นตอนวิธี SCRO I คือการขยายคำค้นโดยใช้อนโทโลยีกฎหมาย เพื่อใช้ค้นคืนในขั้นตอนวิธี SCRO I

2.2 ขั้นตอนวิธี Baseline คือขั้นตอนการจำลองการค้นคืนคำพิพากษาที่ได้รวบรวมไว้

3. ผลการทดลอง

ผู้วิจัยได้ใช้ค่าความหลากหลายของประเด็นทางกฎหมายในการวัดประสิทธิภาพ โดยค่าความหลากหลายของประเด็นทางกฎหมายหาจากจำนวนประเด็นของการค้นคืนแต่ละลำดับ ซึ่งผู้เชี่ยวชาญเป็นผู้พิจารณาคำพิพากษาศาลฎีกาแต่ละลำดับ จำนวนประเด็นทั้งหมดได้จากความสัมพันธ์ที่เกิดขึ้นในอนโทโลยี โดยเปรียบเทียบระหว่างขั้นตอนวิธี SCRO I และขั้นตอนวิธี Baseline โดยคำค้นที่ใช้บนชุดข้อมูลมรดกทั้งหมด 10 คำ เป็นคำที่เกิดขึ้นในอนโทโลยี ใช้ทั้งคำนาม และคำกริยา แสดงดังตารางที่ 17

ตารางที่ 17 ผลการทดลองที่ 2 บนชุดข้อมูลมรดก

คำค้น	จำนวนผลลัพธ์การค้นคืน		จำนวนประเด็นทั้งหมด	ค่าความหลากหลายของประเด็นทางกฎหมาย(DC)	
	SCRO I	Baseline		SCRO I	Baseline
มรดก	11	7	13	0.85	0.54
ผู้จัดการมรดก	13	8	21	0.62	0.38
ทายาท	17	10	20	0.85	0.50
ศาล	2	2	3	0.67	0.67
ผู้ตาย	3	1	4	0.75	0.25
พยาน	3	2	4	0.75	0.50
พินัยกรรม	10	5	13	0.77	0.38
บุตรนอกกฎหมาย	2	1	2	1.00	0.50
ถอน	2	1	3	0.67	0.33
เจ้าหนี้	3	1	4	0.75	0.25
ค่าเฉลี่ย				0.77	0.43

จากตารางที่ 17 ผลการทดลองที่ 2 บนชุดข้อมูลมรดกพบว่าขั้นตอนวิธี SCRO I และขั้นตอนวิธี Baseline สามารถค้นคืนคำพิพากษาศาลฎีกาได้มีประเด็นที่หลากหลายเฉลี่ยเท่ากับ 77% และ 43% ตามลำดับ โดยขั้นตอนวิธี SCRO I สามารถค้นคืนคำที่ได้ความหลากหลายเท่ากับ 100% คือคำว่า “บุตรนอกกฎหมาย” เนื่องจาก ความหลากหลายทั้งหมดของคำนั้นมีประเด็น 2 ประเด็น และคำว่าบุตรนอกกฎหมายเป็นคำที่มีความหมายเฉพาะจึงทำให้สามารถค้นคืนได้ครบทุกความหลากหลาย ส่วนขั้นตอนวิธี Baseline สามารถค้นคืนความหลากหลายได้สูงสุดเท่ากับ 67% ในคำว่า “ศาล” เนื่องจากเป็นคำที่กว้าง และประเด็นเกี่ยวกับศาลมี 3 ประเด็น

ผลการทดลองที่ 2 คำค้นที่ใช้บนชุดข้อมูลตัวเงินทั้งหมด 12 คำ เป็นคำที่เกิดขึ้นในออนไลน์ ใช้ทั้งคำนาม และคำกริยา โดยเปรียบเทียบค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline แสดงดังตารางที่ 18

ตารางที่ 18 ผลการทดลองที่ 2 บนชุดข้อมูลตัวเงิน

คำค้น	จำนวนผลลัพธ์การค้นคืน		จำนวนประเด็นทั้งหมด	ค่าความหลากหลายของประเด็นทางกฎหมาย (DC)	
	SCRO I	Baseline		SCRO I	Baseline
เช็ก	10	7	13	0.77	0.54
ตัวแลกเงิน	3	3	5	0.60	0.60
ตัวสัญญาใช้เงิน	7	7	9	0.78	0.78
ผู้ทรง	4	4	7	0.57	0.57
ธนาคาร	5	4	6	0.83	0.67
หนี้	10	5	13	0.77	0.38
สลักหลัง	2	2	3	0.67	0.67
ดอกเบี้ย	2	1	2	1.00	0.50
โอน	1	1	2	0.50	0.50
ไล่เบี้ย	1	1	2	0.50	0.50
อาวัล	4	3	4	1.00	0.75
ส่งจ่าย	1	1	2	0.50	0.50
ค่าเฉลี่ย				0.71	0.58

จากตารางที่ 18 ผลการทดลองที่ 2 บนชุดข้อมูลตัวเงินพบว่าขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline สามารถค้นคืนคำพิพากษาศาลฎีกาได้มีความหลากหลายของประเด็นทางกฎหมายเท่ากับ 71% และ 58% ตามลำดับ โดยขั้นตอนวิธี SCRO สามารถค้นคืนได้ความหลากหลายสูงสุดเท่ากับ 100% ในคำว่า “ดอกเบี้ย” และ “อาวัล” เนื่องจากค่าความหลากหลายของประเด็นทางกฎหมายทั้งหมดมีค่าน้อยและค่าทั้ง 2 เป็นคำที่มีความหมายไม่กว้าง เมื่อใช้ความสัมพันธ์ที่ได้จากออนโทโลยีมาค้นคืน ทำให้พบคำพิพากษาศาลฎีกาครอบคลุมความหลากหลายส่วนขั้นตอนวิธี Baseline สามารถค้นคืนได้ความหลากหลายสูงสุดเท่ากับ 78% ของคำว่า “ตัวสัญญาใช้เงิน” เนื่องจากกฎหมายเกี่ยวกับตัวสัญญาใช้เงินมีน้อย เมื่อนำคำหลักไปค้นคืนจึงสามารถค้นคำพิพากษาที่หลากหลายออกมาได้

เมื่อนำคำค้นที่เกี่ยวข้องกับคดีมรดกมาพิจารณาค่าความถี่ของคำ ลักษณะของคำ เปรียบเทียบกับค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline โดยเรียงลำดับความถี่ของคำจากมากไปน้อยแสดงดังตารางที่ 19

ตารางที่ 19 การทดลองที่ 2 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ และค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี Baseline และ SCRO I บนชุดข้อมูลมรดก

คำ	ความถี่ ของคำ	ลักษณะ ของคำ	ค่าความหลากหลายของประเด็นทางกฎหมาย	
			Baseline	SCRO I
มรดก	1201	นาม	0.54	0.85
ผู้ตาย	516	นาม	0.25	0.75
ผู้จัดการมรดก	364	นาม	0.38	0.62
พินัยกรรม	347	นาม	0.38	0.77
ศาล	345	นาม	0.67	0.67
ทายาท	340	นาม	0.50	0.85
พยาน	62	นาม	0.50	0.75
บุตรนอกกฎหมาย	8	นาม	0.50	1.00
ถอน	115	กริยา	0.33	0.67
เจ้าหนี้	69	กริยา	0.25	0.75

จากตารางที่ 19 ค่าความถี่ของคำเรียงจากมากไปน้อย และแบ่งกลุ่มตามลักษณะของคำ โดยในการทดลองที่ 2 บนชุดข้อมูลมรดกประเด็นต่างๆ ของกฎหมายทั้งหมดได้จากความสัมพันธ์ทั้งหมดที่เชื่อมระหว่างโนดที่มีความสัมพันธ์กัน ซึ่งคำที่มีค่าความถี่สูง ส่งผลให้ประเด็นต่างๆ ของกฎหมายมีค่าสูง จากผลการทดลอง ค่าความถี่ของคำสูงสุดคือคำว่า “มรดก” ซึ่งขั้นตอนวิธี SCRO I ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.85 และขั้นตอนวิธี Baseline ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.54

เมื่อนำคำค้นที่เกี่ยวข้องกับคดีตัวเงินมาพิจารณาคำความถี่ของคำ ลักษณะของคำ เปรียบเทียบกับค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO I และขั้นตอนวิธี Baseline โดยเรียงลำดับความถี่ของคำจากมากไปน้อยแสดงดังตารางที่ 20

ตารางที่ 20 การทดลองที่ 2 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ และค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO I และ Baseline บนชุดข้อมูลตัวเงิน

คำ	ความถี่ ของคำ	ลักษณะ ของคำ	ค่าความหลากหลายของประเด็นทางกฎหมาย	
			Baseline	SCRO I
เช็ค	1720	นาม	0.54	0.77
ดอกเบี้ย	632	นาม	0.50	1.00
ธนาคาร	602	นาม	0.67	0.83
ตัวสัญญาใช้เงิน	122	นาม	0.78	0.78
ผู้ทรง	104	นาม	0.57	0.57
ตัวแลกเงิน	55	นาม	0.60	0.60
หนี้	2573	กริยา	0.38	0.77
ส่งจ่าย	335	กริยา	0.50	0.50
โอน	265	กริยา	0.50	0.50
สลักหลัง	49	กริยา	0.67	0.67
ไต่เบื้อ	47	กริยา	0.50	0.50
อาวัล	35	กริยา	0.75	1.00

จากตารางที่ 20 ค่าความถี่ของคำเรียงจากมากไปน้อย และแบ่งกลุ่มตามลักษณะของคำ โดยในการทดลองที่ 2 บนชุดข้อมูลตัวเงินประเด็นต่างๆ ของกฎหมายทั้งหมดได้จากความสัมพันธ์ทั้งหมดที่เชื่อมระหว่างโนดที่มีความสัมพันธ์กัน ซึ่งคำที่มีค่าความถี่สูง ส่งผลให้ประเด็นต่างๆ ของกฎหมายมีค่าสูง จากผลการทดลอง ค่าความถี่ของคำสูงสุดคือคำว่า “หนี้” ซึ่งขั้นตอนวิธี SCRO I ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.77 และขั้นตอนวิธี Baseline ค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 0.38

การใช้โครงสร้างของออนโทโลยีของการทดลองที่ 2 สามารถเพิ่มประเด็นความหลากหลายของกฎหมายจากคำพิพากษาศาลฎีกาที่ค้นคืนออกมาได้ เนื่องจากความสัมพันธ์ในโครงสร้างของออนโทโลยีแสดงถึงประเด็นต่างๆ ที่มีความหลากหลาย เมื่อทำการขยายคำค้นขั้นตอนวิธี SCRO I จะแบ่งกลุ่มของคำค้นตามความสัมพันธ์ที่เชื่อมโยงในแต่ละโนด และนำแต่ละกลุ่มคำซึ่งแสดงถึงแต่ละประเด็น เข้าไปค้นคืนคำพิพากษาศาลฎีกา ทำให้ประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายเพิ่มมากขึ้น แต่ค่าในส่วนของความสัมพันธ์ประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายจะมีค่าน้อยกว่าขั้นตอนวิธี SCRO II เนื่องจากไม่มีการขยายไปยังออนโทโลยีระดับที่ 2

การทดลองที่ 3

วัตถุประสงค์ของการทดลองนี้เพื่อเปรียบเทียบประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธีที่ผู้วิจัยได้เสนอคือขั้นตอนวิธี SCRO กับขั้นตอนวิธีที่ผู้วิจัยได้จำลองระบบสืบค้นในปัจจุบันคือขั้นตอนวิธี Baseline เพื่อแสดงประสิทธิภาพของระหว่างการขยายคำค้นแบบกลุ่มเดียว และแบบกลุ่มย่อย

1. ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองรวบรวมจากคำพิพากษาศาลฎีกาที่เกี่ยวข้องกับคดีตัวเงิน มาตรา 898-1001 จำนวน 200 คำพิพากษา

ออนโทโลยีที่ใช้คือออนโทโลยีตัวเงิน ผู้เชี่ยวชาญสร้างจากคำพิพากษาศาลฎีกา และประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน มาตรา 898 ถึง มาตรา 1001 และออนโทโลยีมรดก สร้างจากคำพิพากษาศาลฎีกาและประมวลกฎหมายแพ่งและพาณิชย์บรรพที่ 6 มรดก มาตรา 1599 ถึงมาตรา 1754

2. ผลการทดลอง

ในการทดลองข้อมูลที่ใช้คือคำพิพากษาศาลฎีกาบรรพที่ 3 เอกเทศสัญญา ลักษณะ 21 ว่าด้วยเรื่องตัวเงิน มาตรา 898 ถึงมาตรา 1001 ตั้งแต่ปี พ.ศ. 2542 ถึง พ.ศ. 2552 จำนวน 200 คำ

พิพาทฯ ขั้นตอนที่ใช้คือขั้นตอนวิธี SCRO โดยใช้คำค้น 10 คำค้นในการทดลอง ซึ่งเป็นคำที่เกิดในออนโทโลยี ซึ่งเมื่อได้ขยายคำค้นแล้วได้กลุ่มของคำค้น แสดงในตารางที่ 21

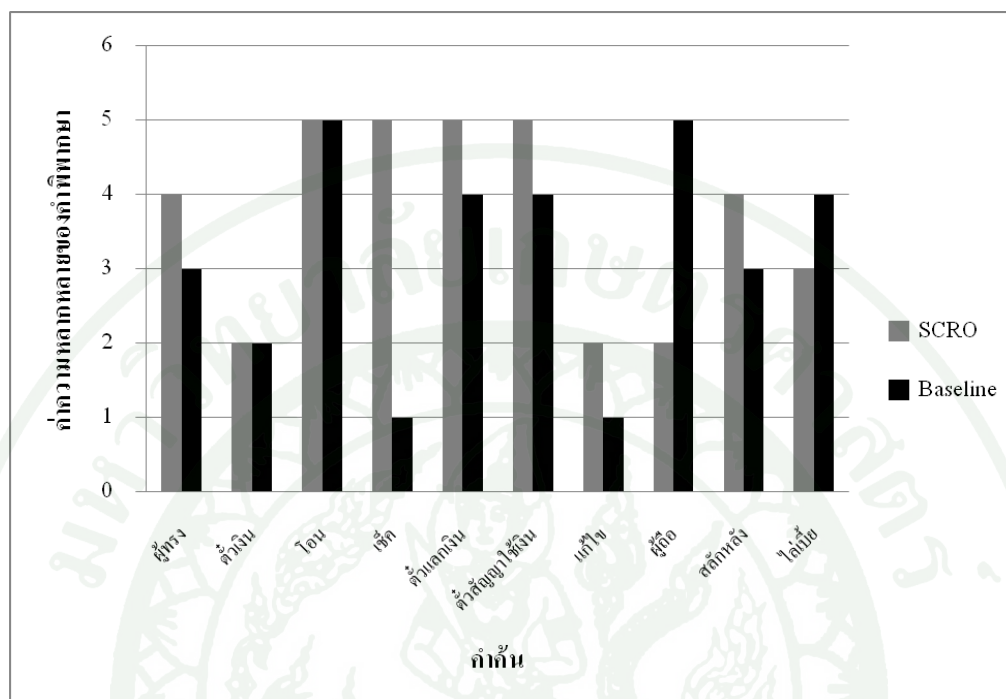
ตารางที่ 21 คำค้นจากผู้ใช้และกลุ่มคำค้นที่เพิ่มขึ้นจากขั้นตอนวิธี SCRO

คำค้นหลักจากผู้ใช้	กลุ่มคำค้นที่เพิ่มขึ้นจากขั้นตอนวิธี SCRO
ผู้ทรง	ผู้รับเงิน ผู้ถือ ไล่เบี้ย วันออกเช็ค สลากหลัง ตัวเงิน เช็ค ตัวแลกเงิน ตัวสัญญาใช้เงิน
ตัวเงิน	เช็ค ตัวแลกเงิน ตัวสัญญาใช้เงิน โอน ผู้ทรง ผู้รับเงิน ผู้ถือ
โอน	ส่งมอบ สลากหลัง ตัวเงิน เช็ค ตัวแลกเงิน ตัวสัญญาใช้เงิน
เช็ค	ตัวเงิน ธนาคาร จดคร่อมทั่วไป จดคร่อมเฉพาะ ขาดลด แก้ไข ปลอมลายมือ มิได้ประทับตรา
ตัวแลกเงิน	ตัวเงิน ผู้ส่งจ่าย ผู้จ่าย
ตัวสัญญาใช้เงิน	ตัวเงิน ใช้เงินเมื่อทวงถาม
แก้ไข	ปลอมลายมือ มิได้ประทับตรา เช็ค ตัวเงิน ตัวแลกเงิน ตัวสัญญาใช้เงิน
ผู้ถือ	ผู้ทรง ไล่เบี้ย วันออกเช็ค สลากหลัง
สลากหลัง	ผู้ทรง ผู้ถือ ผู้รับเงิน โอน ส่งมอบ อาวัล
ไล่เบี้ย	ผู้ทรง ผู้ถือ ผู้รับเงิน

จากตารางที่ 21 ออนโทโลยีขยายคำค้นใหม่จากคำค้นหลักเพื่อใช้ค้นคืนคำพิพาทฯศาลฎีกา ได้กลุ่มของคำค้นใหม่ 1 กลุ่มในแต่ละคำค้น เช่น คำหลักคือ “ผู้ทรง” กลุ่มคำที่ขยายได้คือ “ผู้รับเงิน” “ผู้ถือ” “ไล่เบี้ย” “วันออกเช็ค” “สลากหลัง” “ตัวเงิน” “เช็ค” “ตัวแลกเงิน” และ “ตัวสัญญาใช้เงิน”

คำที่ใช้ในการทดลองมีลักษณะของที่เป็นคำนามคือ “ผู้ทรง” “ตัวเงิน” “เช็ค” “ตัวแลกเงิน” “ผู้ถือ” และ “ตัวสัญญาใช้เงิน” ส่วนลักษณะของที่เป็นคำกริยา คือ “โอน” “แก้ไข” “สลากหลัง” และ “ไล่เบี้ย” ซึ่งประเด็นของกฎหมายที่มีความหลากหลายจะปรากฏในส่วนของคำกริยา ส่วนคำนามจะเป็นส่วนประกอบของการเกิดประเด็นต่างๆ

ผลการทดลองหลังจากนำกลุ่มคำค้นมาใช้ค้นคืนคำพิพากษาศาลฎีกา แสดงดังภาพที่ 27



ภาพที่ 27 การค้นคืนคำพิพากษาที่แตกต่างกันในแต่ละกลุ่มคำค้นของขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline

จากภาพที่ 27 การค้นคืนคำพิพากษาศาลฎีกาโดยวิธี SCRO จะได้คำพิพากษาศาลฎีกาที่มีความหลากหลายในแง่ของประเด็นทางกฎหมาย มากกว่าวิธีการใช้ขั้นตอนวิธี Baseline ในการค้นคืนคำพิพากษาศาลฎีกา

ผู้วิจัยได้ทำการหาค่าความถี่ของคำที่ใช้ในการทดลอง เพื่อพิจารณาค่าความถี่ว่ามีผลต่อการค้นคืนค่าความหลากหลายของประเด็นทางกฎหมายหรือไม่ โดยเปรียบเทียบค่าความถี่ของแต่ละคำลักษณะของคำ และค่าความหลากหลายของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline แสดงดังตารางที่ 22

ตารางที่ 22 การทดลองที่ 3 เปรียบเทียบค่าความถี่ของคำ ลักษณะของคำ และค่าความหลากหลาย
ของประเด็นทางกฎหมายของขั้นตอนวิธี SCRO และ Baseline บนชุดข้อมูลตัวเงิน

คำค้น	ความถี่ ของคำ	ลักษณะ ของคำ	ค่าความหลากหลายของประเด็นทางกฎหมาย	
			SCRO	Baseline
เช็ค	1720	นาม	1.0	0.2
ตัวสัญญาใช้เงิน	122	นาม	1.0	0.8
ผู้ทรง	104	นาม	0.8	0.6
ผู้ถือ	77	นาม	0.4	1.0
ตัวแลกเงิน	55	นาม	1.0	0.8
ตัวเงิน	16	นาม	0.4	0.4
โอน	265	กริยา	1.0	1.0
สลักหลัง	49	กริยา	0.8	0.6
ไต่เบี่ย	47	กริยา	0.6	0.8
แก้ไข	34	กริยา	0.4	0.2

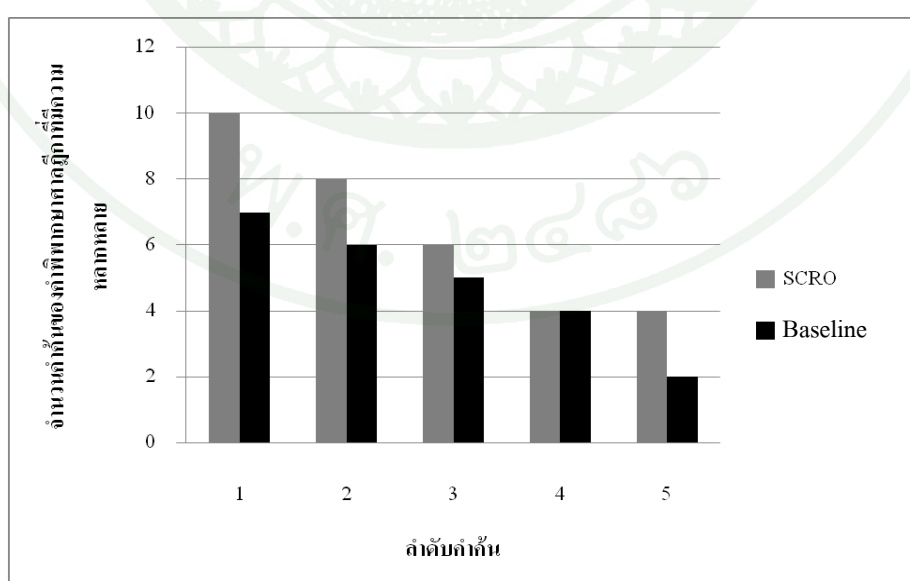
จากตารางที่ 22 ค่าความถี่ของคำเรียงจากมากไปน้อย และแบ่งกลุ่มตามลักษณะของคำ โดยในการทดลองที่ 1 ประเด็นต่างๆ ของกฎหมายทั้งหมด 5 ประเด็น ซึ่งคำที่มีค่าความถี่สูง ส่งผลให้ประเด็นต่างๆ ของกฎหมายมีค่าสูง ทำให้ขั้นตอนวิธี SCRO สามารถค้นคืนได้ค่าความหลากหลายของประเด็นทางกฎหมายสูง เนื่องจากขั้นตอนวิธี SCRO ใช้กลุ่มของคำซึ่งมีค่าหลากหลายๆ ช่วยในการค้นคืนคำพิพากษาศาลฎีกา ส่วนขั้นตอนวิธี Baseline จะให้ค่าความหลากหลายของประเด็นทางกฎหมายต่ำ เมื่อค่าความถี่นั้นมีค่าสูงสุดและต่ำสุด โดยผลการทดลองพบว่าคำว่า “เช็ค” มีค่าความถี่สูงสุด ขั้นตอนวิธี SCRO สามารถให้ค่าความหลากหลายของประเด็นทางกฎหมายเท่ากับ 1.0 และขั้นตอนวิธี Baseline ให้ค่าความถี่เท่ากับ 0.2

เมื่อนำกลุ่มของคำค้นใหม่ มาใช้ในการค้นคืนคำพิพากษาศาลฎีกาที่ได้เก็บรวบรวมไว้ โดยค้นคืนออกมาทั้งหมด 5 ลำดับ และให้ผู้เชี่ยวชาญเป็นผู้วัดประสิทธิภาพของคำพิพากษาศาลฎีกา โดยพิจารณาคำพิพากษาแต่ละลำดับว่าตรงกับความต้องการของผู้ใช้และประเด็นทางกฎหมายมีความหลากหลายระหว่างคำพิพากษาศาลฎีกาหรือไม่ โดยเปรียบเทียบขั้นตอนวิธี SCRO กับขั้นตอนค้นคืนคำพิพากษาศาลฎีกาที่ผู้วิจัยได้จำลองขึ้น แสดงในตารางที่ 23

ตารางที่ 23 ผลการประเมินจากผู้เชี่ยวชาญ

ลำดับเอกสารที่ถูกต้อง และมีความหลากหลาย	ผลการประเมินจากผู้เชี่ยวชาญ (%)	
	SCRO	Baseline
1 อันดับแรก	100	70
2 อันดับแรก	80	60
3 อันดับแรก	60	50
4 อันดับแรก	40	40
5 อันดับแรก	40	20

จากตารางที่ 23 ลำดับคำพินิจภาษาบาลีที่ถูกต้องและมีความหลากหลาย พิจารณาจากผลลัพธ์ของแต่ละคำค้น ว่าลำดับแต่ละลำดับมีประเด็นที่ตรงกับคำค้นที่ผู้ใช้ป้อนเข้ามา หรือไม่ เช่น 1 อันดับแรกหมายถึง ขั้นตอนสามารถค้นคืนคำพินิจภาษาบาลีลำดับที่ 1 ได้ประเด็นตรงกับคำค้นที่ใช้ 2 อันดับแรกหมายถึง ขั้นตอนสามารถค้นคืนคำพินิจภาษาบาลีลำดับที่ 1 และ 2 ได้ตรงกับคำค้น และทั้ง 2 ลำดับมีประเด็นที่ไม่เหมือนกัน จากผลการทดลองแสดงให้เห็นว่าขั้นตอนวิธี SCRO สามารถค้นคืนคำพินิจภาษา 1 อันดับแรกได้ประเด็นที่ตรงกับคำค้นของผู้ใช้เท่ากับ 100% ส่วนขั้นตอนวิธี Baseline สามารถค้นคืนได้ตรงกับคำค้นของผู้ใช้เท่ากับ 70% และทั้ง 2 ขั้นตอนประสิทธิภาพในการค้นคืนให้ได้คำพินิจภาษาบาลีมีประสิทธิภาพที่ลดลง แสดงในภาพที่ 28



ภาพที่ 28 เปรียบเทียบประสิทธิภาพระหว่างขั้นตอนวิธี SCRO และขั้นตอนวิธี Baseline

จากภาพที่ 28 ลำดับคำพิพากษาที่มีความหลากหลายของประเด็นทางกฎหมายเปรียบเทียบกับจำนวนคำค้น เมื่อพิจารณาที่ลำดับคำพิพากษาที่มีความหลากหลาย 1 อันดับ คำค้นของขั้นตอนวิธี Baseline มี 7 คำค้นที่สามารถค้นคืนได้ตรงกับประเด็นทางกฎหมาย ส่วนขั้นตอนวิธี SCRO มี 10 คำ ที่ค้นคืนได้ตรงกับประเด็นทางกฎหมาย ถ้าพิจารณาลำดับคำพิพากษาศาลฎีกาที่มีความหลากหลาย 2 อันดับ ขั้นตอนวิธี Baseline มี 6 คำค้นที่สามารถค้นคืนได้มีความหลากหลาย 2 อันดับ ส่วนขั้นตอนวิธี SCRO มี 8 คำที่ค้นคืนได้มีความหลากหลาย 2 อันดับ เมื่อลำดับคำพิพากษาที่มีความหลากหลายเพิ่มขึ้น พบว่าจำนวนคำค้นที่ค้นคืนได้มีความหลากหลายแต่ละลำดับมีจำนวนลดลง เนื่องจากประเด็นทางกฎหมายลดลง และเมื่อค้นคืนออกมาจะซ้ำกับประเด็นแรก แต่ขั้นตอนวิธี SCRO มีจำนวนคำค้นที่สามารถค้นคืนได้มีความหลากหลายของแต่ละลำดับมีมากกว่าขั้นตอนวิธี Baseline

เมื่อเปรียบเทียบลำดับค่าความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธี SCRO, SCRO I และ SCRO II พบว่าจำนวนค่าความหลากหลายของขั้นตอนวิธี SCRO จะลดลงเมื่อลำดับที่ต้องการความหลากหลายของประเด็นทางกฎหมายเพิ่มขึ้น แต่จำนวนค่าความหลากหลายของขั้นตอนวิธี SCRO I และ SCRO II ไม่ได้ขึ้นอยู่กับลำดับที่เพิ่มขึ้น แต่ขึ้นอยู่กับความสัมพันธ์ของออนโทโลยีที่ขยายออกมาเป็นกลุ่มของคำค้น ดังนั้นในการขยายค่าแบบกลุ่มเดียวจึงทำให้ไม่สามารถหาคำพิพากษาศาลฎีกาที่มีความหลากหลายได้

การขยายคำค้นแบบกลุ่มเดียว ไม่สามารถค้นคืนคำพิพากษาศาลฎีกาให้มีประเด็นที่มีความหลากหลายได้ เนื่องจากในการค้นคืนคำพิพากษาศาลฎีกา ขั้นตอนวิธีจะนำคำทั้งหมดในกลุ่มคำค้นที่ขยายออกมาได้ ไปค้นคืนคำพิพากษาศาลฎีกาพร้อมกัน โดยเชื่อมด้วยคำว่า “และ” และเรียงลำดับของผลลัพธ์ของการค้นคืนตามน้ำหนักของค่า *TFIDF* จากมากไปน้อยซึ่งไม่สามารถแยกถึงประเด็นต่างๆ ได้ ส่วนการค้นคืนโดยใช้การแบ่งกลุ่มคำค้นออกเป็นกลุ่มย่อยๆ สามารถค้นคืนคำพิพากษาศาลฎีกาให้มีประเด็นทางกฎหมายที่มีความหลากหลายได้ เนื่องจากการค้นคืนคำพิพากษาศาลฎีกาจะนำกลุ่มคำค้นย่อย ซึ่งแสดงถึงประเด็นต่างๆ ของกฎหมายจากความสัมพันธ์ของออนโทโลยี ไปค้นคืนคำพิพากษาศาลฎีกา ซึ่งแสดงถึงคำพิพากษาศาลฎีกาตามกลุ่มคำค้นย่อย และขั้นตอนวิธี SCRO จะเลือกคำพิพากษาศาลฎีกาลำดับที่ 1 ออกมาในแต่ละกลุ่ม โดยถ้ามีคำพิพากษาศาลฎีกาซ้ำกันจะเลือกคำพิพากษาศาลฎีกาลำดับถัดไปในกลุ่มที่มีค่าน้ำหนักน้อย ดังนั้น ผลลัพธ์ของการค้นคืนจึงได้คำพิพากษาศาลฎีกาในแต่ละประเด็นที่มีความหลากหลาย

วิจารณ์

การทดลองที่ 1 ใช้โครงสร้างของออนโทโลยีมีระดับ 2 ชั้น โดยชั้นที่ 1 คือโครงสร้างความสัมพันธ์หลักของประเด็นต่างๆ และชั้นที่ 2 คือส่วนขยายของคำในโนดและความสัมพันธ์ในออนโทโลยี ส่วนการทดลองที่ 2 ใช้โครงสร้างของออนโทโลยีมี 1 ระดับชั้นคือโครงสร้างความสัมพันธ์หลักของประเด็นต่างๆ เท่านั้น ซึ่งกลุ่มคำค้นที่ขยายได้จากทั้ง 2 การทดลองนี้ เป็นกลุ่มคำค้นย่อย ที่แสดงถึงประเด็นต่างๆ ของกฎหมาย และการทดลองที่ 3 ใช้โครงสร้างของออนโทโลยีเป็นโครงสร้างต้นไม้ ซึ่งความสัมพันธ์แสดงถึงความสัมพันธ์ระหว่างคำ ไม่ได้แสดงถึงประเด็นต่างๆ และกลุ่มของคำค้นที่ขยายได้ มีเพียง 1 กลุ่มคำค้น

โครงสร้างของออนโทโลยีของการทดลองที่ 1 และ 2 สามารถแสดงถึงประเด็นต่างๆ ของกฎหมายได้ โดยอาศัยความสัมพันธ์ที่เชื่อมระหว่างโนด 2 โหนดซึ่งส่งผลให้ขั้นตอนในการขยายคำค้น สามารถขยายคำค้นออกเป็นกลุ่มย่อยๆ ได้ตามความสัมพันธ์ของออนโทโลยี เมื่อนำกลุ่มคำค้นย่อยแต่ละกลุ่มค้นคืนคำพิพากษาศาลฎีกา สามารถค้นคืนคำพิพากษาศาลฎีกาในแต่ละประเด็นได้ตามกลุ่มของคำค้น แต่ในการทดลองที่ 3 กลุ่มของคำค้นมีเพียงกลุ่มเดียว เนื่องจากโครงสร้างของออนโทโลยีไม่ได้แสดงถึงประเด็นต่างๆ ทางกฎหมาย ทำให้เมื่อนำกลุ่มคำค้นที่ได้ไปค้นคืนเอกสาร จึงได้ผลลัพธ์ของการค้นคืนที่มีค่าความหลากหลายของประเด็นทางกฎหมายน้อยกว่าการทดลองที่ 1 และ 2

กลุ่มคำค้นย่อยที่ได้จากโครงสร้างของออนโทโลยีของการทดลองที่ 1 และ 2 มีผลต่อค่าความหลากหลายของประเด็นทางกฎหมาย เนื่องจากแต่ละกลุ่มคำค้นย่อยแสดงถึงแต่ละประเด็นที่ได้จากโครงสร้างความสัมพันธ์ของออนโทโลยี เมื่อนำไปค้นคืนคำพิพากษาศาลฎีกาทำให้ได้คำพิพากษาศาลฎีกาตามประเด็นของกลุ่มคำค้นย่อย และเมื่อขั้นตอนวิธีเลือกผลลัพธ์ลำดับที่ 1 ของแต่ละกลุ่มผลลัพธ์ จะได้คำพิพากษาศาลฎีกาที่มีความหลากหลายทางประเด็นของกฎหมาย แต่กลุ่มคำค้นที่ขยายออกเป็นกลุ่มเดียวของการทดลองที่ 3 ไม่สามารถแยกประเด็นต่างๆ ของกฎหมายได้ เนื่องจากในการค้นคืนคำพิพากษาศาลฎีกา ใช้คำทั้งหมดจากกลุ่มคำค้นค้นคืนคำพิพากษาศาลฎีกาพร้อมกัน โดยเชื่อมด้วยคำว่า “และ” ซึ่งในคำพิพากษาศาลฎีกาจะมีโอกาสที่มีทุกคำบรรจุอยู่น้อยมาก ทำให้ผลลัพธ์ที่ได้เกิดจากการเรียงลำดับตามค่าน้ำหนักของค่า *TFIDF* เท่านั้น

ระดับของออนโทโลยีของการทดลองที่ 3 มี 2 ระดับ โดยระดับที่ 1 แสดงถึงโครงสร้าง ความสัมพันธ์ของแต่ละประเด็น โนคและความสัมพันธ์สามารถขยายไปสู่ออนโทโลยีระดับ 2 ได้ ซึ่งโนคขยายคำที่มีความหมายเหมือนกันและความสัมพันธ์ขยายคำที่มีความสัมพันธ์กัน ทำให้ สามารถขยายคำค้นได้มีประสิทธิภาพดีกว่าโครงสร้างออนโทโลยีที่มีเพียงระดับเดียว ซึ่งแสดงเพียง โครงสร้างความสัมพันธ์ของประเด็นหลักเท่านั้น

งานวิจัยนี้ได้ศึกษาการใช้ออนโทโลยีสำหรับการค้นคืนคำพิพากษาศาลฎีกา ซึ่งในปัจจุบัน (พ.ศ. 2553) ออนโทโลยีเกี่ยวกับกฎหมายของไทย ยังไม่มีการวิจัยและสร้างขึ้นมา ดังนั้น ออนโทโลยีที่ใช้ในงานวิจัยนี้ จึงได้มีการออกแบบโครงสร้างของออนโทโลยี โดยให้ผู้เชี่ยวชาญ เป็นผู้กำหนดความสัมพันธ์ของคำในโนคต่างๆ โดยผู้เชี่ยวชาญจะพิจารณาคำในโนคและ ความสัมพันธ์ต่างๆ จากประมวลกฎหมายแพ่งและพาณิชย์ และคำพิพากษาศาลฎีกา โดยการเลือก คำพิจารณาจากคำที่มีความสัมพันธ์ทางด้านประเด็นของกฎหมาย เช่น ประเด็นมรดกตกทอดให้แก่ ทายาท คำที่เลือกออกมาสามารถสร้างโนคที่มีความสัมพันธ์ระหว่างคำว่า “มรดก” และ “ทายาท” และความสัมพันธ์ที่ได้คือคำว่า “ตกทอด”

ผลการทดลองกับระบบสืบค้นคำพิพากษาและคำสั่งคำร้องศาลฎีกา โดยเปรียบเทียบ ระหว่างขั้นตอนวิธี Deka II และ Deka I พบว่า ขั้นตอนวิธี Deka II คือขั้นตอนที่นำกลุ่มคำค้นย่อยที่ ได้จากการขยายคำค้นโดยออนโทโลยี สามารถค้นคืนได้ค่าความหลากหลายของประเด็นทาง กฎหมายมากกว่าขั้นตอนวิธี Deka I แสดงให้เห็นว่าการขยายคำค้นโดยใช้ออนโทโลยีจากขั้นตอน วิธี SCRO II สามารถเพิ่มประสิทธิภาพของค่าความหลากหลายของประเด็นทางกฎหมายให้กับ ระบบจริงได้

สรุปและข้อเสนอแนะ

สรุป

วัตถุประสงค์ของงานวิจัยเพื่อเสนอขั้นตอนวิธีในการค้นคืนคำพิพากษาศาลฎีกาให้มี ประเด็นทางกฎหมายที่มีความหลากหลาย โดยผู้วิจัยสร้างขั้นตอนวิธี SCRO คือการใช้ออนโทโลยี โดยผู้เชี่ยวชาญสร้างจากประมวลกฎหมายแพ่งและพาณิชย์คดีมรดกและตัวเงิน ขยายคำค้นจาก ความสัมพันธ์ของออนโทโลยี และนำกลุ่มของคำค้นใช้ในการค้นคืนคำพิพากษาศาลฎีกา โดยได้ทำ การทดลองกับคำพิพากษาศาลฎีกา ในคดีมรดกและตัวเงิน โดยแบ่งการทดลองออกเป็น 3 การ ทดลองดังต่อไปนี้

ผลการทดลองที่ 1 ใช้โครงสร้างของออนโทโลยีให้มี 2 ระดับ เพื่อให้ โหนดและ ความสัมพันธ์ สามารถขยายคำที่มีความสัมพันธ์ หรือเป็นคำที่มีความหมายเหมือนกันได้ โดย ทดลองกับคำพิพากษาศาลฎีกาบนชุดข้อมูลมรดกและตัวเงิน เปรียบเทียบกับขั้นตอนวิธี SCRO I ขั้นตอนวิธี Baseline ผลการทดลองแสดงให้เห็นว่าขั้นตอนวิธี SCRO II สามารถค้นคืนคำพิพากษา ศาลฎีกาที่มีประเด็นทางกฎหมายที่มีความหลากหลายได้ในส่วนคำที่เป็นความสัมพันธ์มากกว่า ขั้นตอนอื่นๆ โดยผลการทดลองที่ 1 บนชุดข้อมูลมรดกและตัวเงิน ขั้นตอนวิธี SCRO II ให้ค่าความ หลากหลายเท่ากับ 72% และ 79% ตามลำดับ

ต่อมาทดลองกับขั้นตอนการค้นคืนคำพิพากษาศาลฎีกาในปัจจุบัน โดยเปรียบเทียบค่า ความหลากหลายของประเด็นทางกฎหมายระหว่างขั้นตอนวิธี Deka I และขั้นตอนวิธี Deka II ผล การทดลองแสดงให้เห็นว่า การขยายคำค้นจากขั้นตอนวิธี SCRO II สามารถเพิ่มค่าความ หลากหลายของประเด็นทางกฎหมายให้กับขั้นตอนวิธี Deka II ได้ โดยการทดลองที่ 1 บนชุดข้อมูล มรดกและตัวเงิน ขั้นตอนวิธี Deka II ให้ค่าความหลากหลายเท่ากับ 0.25 และ 0.26 ตามลำดับ

ผลการทดลองที่ 1 แสดงให้เห็นว่าการขยายคำค้นโดยใช้ออนโทโลยี 2 ระดับสามารถเพิ่ม ประสิทธิภาพของค่าความหลากหลายทางประเด็นของกฎหมายได้มากกว่าการใช้ออนโทโลยี 1 ระดับ เนื่องจากออนโทโลยีระดับที่ 2 จะขยายคำในส่วนที่เป็นคำที่มีความหมายเหมือนกันและคำที่ มีความสัมพันธ์กันเข้าไปยังกลุ่มของคำค้นย่อย ทำให้ได้ประเด็นของกฎหมายเพิ่มมากขึ้น

ผลการทดลองที่ 2 ใช้ออนโทโลยี 1 ระดับ แสดงความความสัมพันธ์ระหว่างคำที่ใช้แทนความหลากหลายของประเด็นทางกฎหมาย โดยผลการทดลองเปรียบเทียบกับวิธีของการค้นคืนโดยใช้ความถี่ แสดงให้เห็นว่าขั้นตอนวิธี SCRO I สามารถค้นคืนคำพิพากษาศาลฎีกาได้มีประเด็นทางกฎหมายที่มีความหลากหลาย แต่ได้ค่าความหลากหลายของประเด็นทางกฎหมายน้อยกว่าขั้นตอนวิธี SCRO II เนื่องจากโครงสร้างของออนโทโลยี 2 ชั้น ในชั้นที่ 1 จะแสดงถึงความสัมพันธ์โดยทั่วไป และในชั้นที่ 2 จะขยายคำจากโนดและความสัมพันธ์ โดยโนดจะขยายคำที่มีความหมายเหมือนกัน ซึ่งแสดงถึงประเด็นต่างๆ ส่วนความสัมพันธ์จะขยายคำในส่วนลักษณะของแต่ละประเด็น

ผลการทดลองที่ 3 มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของลำดับของค่าความหลากหลายของประเด็นทางกฎหมาย โดยขั้นตอนวิธี SCRO จะขยายคำค้นออกมาเพียง 1 กลุ่ม แต่ขั้นตอนวิธี SCRO II จะขยายคำค้นออกเป็นกลุ่มย่อยๆ ผลการทดลองแสดงให้เห็นว่า กลุ่มคำค้นของขั้นตอนวิธี SCRO ไม่สามารถค้นคืนคำพิพากษาศาลฎีกาโดยให้ค่าความหลากหลายของประเด็นทางกฎหมายที่สูงได้ เนื่องจากกลุ่มของคำค้นเป็นกลุ่มเดียวไม่มีการแบ่งประเด็นทางกฎหมายออกเป็นกลุ่มย่อยๆ

ข้อเสนอแนะ

งานวิจัยนี้มุ่งเน้นในการนำออนโทโลยีมาใช้ในการขยายคำค้น เพื่อสร้างกลุ่มของคำค้นให้สามารถค้นคืนคำพิพากษาศาลฎีกาที่มีประเด็นที่มีความหลากหลายได้ แต่การสร้างออนโทโลยีสำหรับกฎหมายแต่ละด้านใช้เวลานาน และต้องอาศัยผู้เชี่ยวชาญเป็นผู้ตรวจสอบ ดังนั้น จึงควรวิจัยเพิ่มทางด้านการสร้างออนโทโลยีกฎหมายโดยอัตโนมัติ เพื่อลดเวลาในการสร้างออนโทโลยี

เอกสารและสิ่งอ้างอิง

- เว็บไซต์ระบบสืบค้นคำพิพากษาและคำสั่งศาลฎีกา. 2007. ระบบสืบค้นคำพิพากษาและคำสั่งศาลฎีกา. แหล่งที่มา: <http://www.deka2007.supremecourt.or.th/> วันที่สืบค้น 1 สิงหาคม 2552
- Alani, H., N. Noy., N. Shah., N. Shadbolt and M. Musen. 2007. Searching Biomedical Ontologies Based on Content. In **10th Intl. Protégé Conference**.
- Ahmad, K. 2005. Building e-Laws Ontology: New Approach. pp. 826–835. In **OTM Workshops**.
- Andreou, A. 2005. Ontologies and Query expansion.
- Benjamins, V. R., J. Contreras, P. Casanovas, M. Ayuso, Becue, M., L. Lemus, and C. Urios. 2006. Ontologies of professional legal knowledge as the basic for intelligent IT support for judges. pp. 359-378. In **Artificial Intelligence and Law**
- Bratsas, C., V. Koutkias, E. Kaimakamis, P. Bamidis, and N. Maglaveras. 2007. Ontology-based Vector Space Model and Fuzzy Query Expansion to Retrieve Knowledge on Medical Computational Problem Solutions, pp. 23-26. In **Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale**.
- Drid, O. and M. Ben Ahmed. 2008. Building an Ontology-Based Framework For Semantic Information Retrieval: Application To Breast Cancer. In **3rd International Conference**.
- Fiaidhi, J., S. Mohammed, J. Jaam and A. Hasnah. 2003. A Standard Framework for Personalization Via Ontology-Based Query Expansion. In **2nd Pakistan Journal of Information and Technology**.

- Guiraud, L. and Sylvie, D. 2005. Updating ontology in the legal domain. In **Proc of the 10th international conference on Artificial intelligence and law.**
- Gruber, T.R. 1993. Towards principles for the design of ontologies used for knowledge sharing. In Guarino, N. and Poli, R., editors, **Formal Ontology in Conceptual Analysis and Knowledge Representation, Deventer, The Netherlands. Kluwer Academic Publishers.**
- Henze, N., P.Dolog, and W. Nejdl. 2004. Reasoning and Ontologies for Personalized E-Learning in the Semantic Web. pp.82-97. In **Educational Technology & Society.**
- Hohfeld, W. Fundamental legal conceptions as applied in legal reasoning. In **Yale University Press, London.**
- Jimeno-Yepes, A., R. Berlanga-Llavori and D. Rebholz-Schuhmann. 2009. Terminological cleansing for improved information retrieval based on ontological terms, pp. 6-14. In **Proceedings of the WSDM '09 Workshop on Exploiting Semantic Annotations in Information Retrieval.**
- Jovic, A., M. Prcela and D. Gamberger. 2007. Ontologies in Medical Knowledge Representation, pp. 535-540. in **Proc. of the 29th Int. Conf. on Information Technology Interfaces.**
- Kobayashi, M. and K. Takeda. 2000. Information Retrieval on the Web. In **ACM Computing Surveys (CSUR).**
- Marie-Francine, M. 2001. Innovative techniques for legal text retrieval, pp. 29-57. In **Artificial Intelligence and Law.**

- Mustafa, J., S. Khan. and K. Latif. 2008. Ontology Based Semantic Information Retrieval. In **4th International IEEE Conference "Intelligent Systems"**.
- Salton, G., A. Wong and C. S. Yang. 1975. A Vector Space Model for Automatic Indexing, pp. 613-620. In **Communications of The ACM**.
- Saravanan, M., B. Ravindran. and S. Raman. 2009. sing an ontological framework. In **Springer Science Business Media B.V**.
- Schweighofer, E. and A. Geist. 2007. Legal Query Expansion using Ontologies and Relevance Feedback, pp. 149-160. In **Dblp**.
- Soonhee, H., J. Youngim, Y. Aesun, and H. C.Kwon. 2006. Building Korean Classifier Ontology Based on Korean WordNet pp. 261-268.
- Tuominen, J., T Kauppinen., K. Viljanen and E. Hyvönen. 2009. Ontology-Based Query Expansion Widget for Information Retrieval. In **Proc. of 5th Workshop on Scripting and Development for the Semantic Web at ESWC**.
- Valente, A. and J. Breuker. 1994. Making ends meet: conceptual models and ontologies in legal problem solving. pp. 1-15. In **Proc. Of the XI Brazillian AI Symposium**.
- Wu, F., G. Wu and X. Fu. 2007. Design and Implementation of Ontology-Based Query Expansion for Information Retrieval, pp. 293-298. In **Springer Boston**.
- Xu, J. and W. Bruce Croft. 2000. Improving the Effectiveness of Information Retrieval with Local Context Analysis, pp. 79-112. In **ACM Transactions on Information Systems**.



ภาคผนวก



ตารางผนวกที่ ก1 คำที่ใช้ในการสร้างออนโทโลยี 260 คำ

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
1	กรรมสิทธิ์	75	86
2	กองมรดก	60	3
3	ก่อน	122	364
4	การสละมรดก	10	0
5	กำหนดมิให้รับมรดก	6	0
6	กำหนด	135	593
7	แก้ไข	12	107
8	ข้อกำหนด	36	25
9	ข้อความ	66	189
10	ขัด	41	72
11	ขายทอดตลาด	8	50
12	ขายสด	1	49
13	จีด	9	40
14	จีดมา	9	29
15	บูด	2	0
16	บูดลบ	2	0
17	คนเดียว	6	23
18	คนล้มละลาย	3	0
19	คนละส่วน	4	4
20	ครอบครอง	103	62
21	ควบคุม	1	17
22	ความรับผิดชอบ	32	118
23	ความรับผิดชอบ	0	8
24	คัดค้าน	331	156
25	ค้างชำระ	5	90
26	คำประกัน	22	216

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
27	คำสั่ง	94	352
28	คำสั่งศาล	16	24
29	คุณสมบัติ	6	6
30	คู่สมรส	15	0
31	เงิน	170	3035
32	เงื่อนไข	13	129
33	เงื่อนไขเวลา	1	4
34	จดทะเบียน	112	146
35	จัดการ	521	105
36	จัดการทรัพย์สิน	12	7
37	จัดการมรดก	461	13
38	จ่ายเงิน	7	259
39	จำนอง	76	219
40	จำหน่าย	19	69
41	เจตนา	48	141
42	เจ้าของ	72	108
43	เจ้ามรดก	142	0
44	เจ้าหนี้	60	384
45	หนี้	9	29
46	โฉนด	72	35
47	ชั้น	117	592
48	ชำระ	112	1630
49	ชำระหนี้	60	890
50	เช็ค	14	1720
51	เช่า	54	262
52	เช่าซื้อ	13	205

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
53	ใช้เงิน	6	289
54	ซื้อขาย	26	164
55	ดอกเบี้ย	13	633
56	ดูแล	14	50
57	เดิม	35	117
58	โดยทำเป็นหนังสือ	0	3
59	โดยพินัยกรรม	10	0
60	โดยลำพัง	4	4
61	ตก	194	518
62	ตกทอด	48	5
63	ตราตำแหน่ง	1	0
64	ต่อหน้า	17	13
65	ตั้ง	194	235
66	ตัดมิให้รับมรดก	6	0
67	ตัวแทน	7	109
68	ตัวแลกเงิน	0	55
69	ตัวสัญญาใช้เงิน	0	122
70	ตา	1603	3714
71	ต่างประเทศ	3	30
72	ต่างมารดา	2	0
73	ตาย	668	39
74	ตีราคา	2	5
75	เดิม	10	42
76	แต่ง	15	32
77	แต่งงาน	3	0

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
78	แต่งตั้ง	12	20
79	ถอดถอน	25	1
80	ถอน	142	199
81	ถึงแก่ความตาย	102	9
82	ถือ	173	598
83	ถูกจำกัดมิให้รับมรดก	6	0
84	ทรัพย์	486	556
85	ทรัพย์มรดก	296	8
86	ทรัพย์สิ้น	105	218
87	ทวงถาม	0	45
88	ทหาร	7	12
89	ทะเบียน	127	173
90	ทายาท	354	17
91	ทายาทโดยธรรม	93	1
92	ทารกในครรภ์	0	0
93	ทำพินัยกรรม	123	0
94	ทำลาย	6	7
95	ทำศพ	14	0
96	ที่ดิน	626	684
97	ธนาคาร	33	602
98	นอกกฎหมาย	8	0
99	น้อง	16	7
100	น้ำ	142	385
101	นิติกรรม	38	66
102	นิติบุคคล	4	55

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
103	บังคับ	144	526
104	บัญญัติ	39	661
105	บัญญัติ	111	495
106	บันทึก	24	60
107	บางส่วน	10	55
108	บ้าน	47	34
109	บำเหน็จ	5	5
110	บิดา	58	3
111	บิดามารดา	17	1
112	บิดามารดาเดียวกัน	7	0
113	บุคคล	65	259
114	บุคคลวิกลจริต	2	0
115	บุญธรรม	15	0
116	บุตร	112	7
117	บุตรนอกกฎหมาย	8	0
118	บุตรบุญธรรม	15	0
119	บุพการี	1	0
120	บุริมสิทธิ	1	3
121	เบี้ยประกันภัย	2	19
122	แบ่ง	179	49
123	แบ่งปัน	48	3
124	แบ่งมรดก	42	0
125	ปฏิเสธ	9	194
126	ประกันชีวิต	2	11
127	ประกันภัย	23	150
128	ประชุม	3	19

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
129	ประนีประนอมยอมความ	19	134
130	ประโยชน์	67	175
131	ปลด	3	17
132	ปลดหนี้	0	3
133	ปลอม	6	66
134	ป็น	54	4
135	ป่า	1	0
136	ปิดบัง	12	0
137	ปิดบังพินัยกรรม	2	0
138	เปลี่ยนแปลง	6	52
139	เปลี่ยนแปลงพินัยกรรม	2	0
140	ผ่อน	6	75
141	ผู้เขียน	3	2
142	ผู้จัดการทำศพ	3	0
143	ผู้จัดการมรดก	396	13
144	ผู้ตาย	505	22
145	ผู้ถือ	10	77
146	ผู้ทรง	5	104
147	ผู้ทำพินัยกรรม	41	0
148	ผู้แทนโดยชอบธรรม	2	0
149	ผู้ปกครอง	1	0
150	ผู้พิทักษ์	1	0
151	ผู้มีส่วนได้เสีย	70	5
152	ผู้เยาว์	11	1
153	ผู้รับ	78	174
154	ผู้รับประกัน	4	35

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
155	ผู้รับประกันภัย	4	33
156	ผู้รับประกัน	6	14
157	ผู้รับประกันกรรม	20	0
158	ผู้รับมรดก	16	0
159	ผู้รับมรดกแทนที่	1	0
160	ผู้สืบสันดาน	17	1
161	พนักงานเจ้าหน้าที่	14	14
162	พนักงานอัยการ	3	16
163	พยาน	57	297
164	พระภิกษุ	15	0
165	พินัยกรรม	414	1
166	พี่	18	6
167	พี่น้องร่วมบิดามารดา	5	0
168	เพิกถอน	64	63
169	แพทย์	2	10
170	ฟ้อง	291	1293
171	ภริยา	54	30
172	ภานี	21	219
173	มรณะ	4	4
174	มรดก	1293	29
175	มอบอำนาจ	1	93
176	มารดา	47	7
177	มิได้รับมรดก	17	0
178	มีเงื่อนไข	4	15
179	มูลนิธิ	3	1
180	โมฆะ	41	60

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
181	โมณะ	41	60
182	ไม่เกินกว่า	3	2
183	ไม่มีมูลหนี้	0	18
184	ไม่สมบูรณ์	8	19
185	ชักย้าย	5	2
186	ย่า	89	348
187	ยาย	15	184
188	แยกกันอยู่	4	1
189	ระบุ	57	299
190	รับ	501	2284
191	รับ	501	2284
192	รับมรดก	124	1
193	รับมรดกแทนที่	7	0
194	รายงาน	4	38
195	เรียกเก็บ	0	212
196	ไร้ความสามารถ	10	0
197	ลงทุน	2	35
198	ลงลายมือชื่อ	43	130
199	ลบ	27	78
200	ล้มละลาย	7	178
201	ลายนิ้วมือ	1	2
202	ลายพิมพ์นิ้วมือ	6	1
203	ลายมือ	67	166
204	ลายมือชื่อ	62	165
205	ลำดับ	14	19
206	ลำพัง	6	7

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
207	ลุง	3	3
208	ลุง ป้า น้า อา	0	0
209	ลูกหนี้	32	317
210	ไล่	9	50
211	ไล่เบี้ย	1	48
212	วัด	48	81
213	วาจา	1	6
214	วิกตจริต	2	0
215	ศพ	23	9
216	ศาล	350	1590
217	สติ	4	1
218	สมรส	63	33
219	สละ	29	31
220	สละมรดก	16	0
221	สลักหลัง	0	49
222	ส่วนแบ่ง	19	2
223	สั่งจ่าย	1	335
224	สังหาริมทรัพย์	10	20
225	สัญญา	148	1748
226	สัญญาประกันชีวิต	2	2
227	สัญญาประนีประนอม	18	117
228	สัดส่วน	1	3
229	สามี	45	22
230	สิ่งปลูกสร้าง	16	25
231	สิทธิ	447	761
232	สินเดิม	4	0

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
233	สินสมรส	17	23
234	สินไหมทดแทน	5	40
235	สีก	1	0
236	สืบหา	1	0
237	สูญหาย	1	33
238	เสียงข้างมาก	2	0
239	หนังสือ	73	335
240	หน้าที่	117	268
241	หนี้	223	2580
242	หนี้	223	2580
243	หย่า	7	7
244	หลอกลวง	4	23
245	หลัง	80	255
246	หลายคน	5	5
247	หุ้น	19	211
248	ให้เช่า	3	21
249	อยู่กิน	22	3
250	อสังหาริมทรัพย์	8	18
251	อา	280	1240
252	อายุ	89	173
253	อายุความ	82	163
254	อาวัล	0	35
255	อำนาจ	105	463
256	อำเภอ	38	31
257	เอกสาร	70	379
258	เอกสารฝ่ายเมือง	19	0

ตารางผนวกที่ ก1 (ต่อ)

ลำดับ	คำ	ความถี่	
		มรดก	ตัวเงิน
259	โอน	130	266
260	โอนสิทธิ	3	40

ตารางผนวกที่ ก2 คำในโนดของออนโทโลยีมรดกที่ใช้ในขั้นตอนวิธี SCRO II

ลำดับ	ออนโทโลยีระดับ 1	ออนโทโลยีระดับ 2
1	มรดก	ทรัพย์สินมรดก กองมรดก ทรัพย์สิน สิทธิ หน้าที่ ความรับผิดชอบ
2	ทายาท	ทายาทโดยธรรม ผู้รับพินัยกรรม ผู้รับมรดก ผู้เยาว์ บุคคลวิกลจริต พระภิกษุ ผู้มีส่วนได้เสีย
3	ผู้สืบสันดาน	บุตรนอกกฎหมาย บุตรบุญธรรม ภริยา
4	ผู้จัดการมรดก	ผู้จัดการทำศพ
5	ผู้ตาย	เจ้ามรดก

ตารางผนวกที่ ก2 (ต่อ)

ลำดับ	ออนโทโลยีระดับ 1	ออนโทโลยีระดับ 2
6	ทฤษฎีโดยธรรมชาติ	ผู้สืบสันดาน บิดามารดา พี่น้องร่วมบิดามารดา ปู่ ย่า ตา ยาย ลุง ป้า น้า อา ทารกในครรภ์ คู่สมรส

ตารางผนวกที่ ก3 คำในความสัมพันธ์ของออนโทโลยีมรดกที่ใช้ในขั้นตอนวิธี SCRO II

ลำดับ	ออนโทโลยีระดับ 1	ออนโทโลยีระดับ 2
1	คำจำกัดความให้รับมรดก	ปิดบัง ชักย้าย ปลอม ทำลาย ปิดบังพินัยกรรม เปลี่ยนแปลงพินัยกรรม ฉ้อฉล
2	ตัดมรดกให้รับมรดก	โดยพินัยกรรม โดยทำเป็นหนังสือ
3	การสละมรดก	บางส่วน มีเงื่อนไข สัญญาประนีประนอม
4	ตั้ง	คนเดียว หลายคน

ตารางผนวกที่ ก3 (ต่อ)

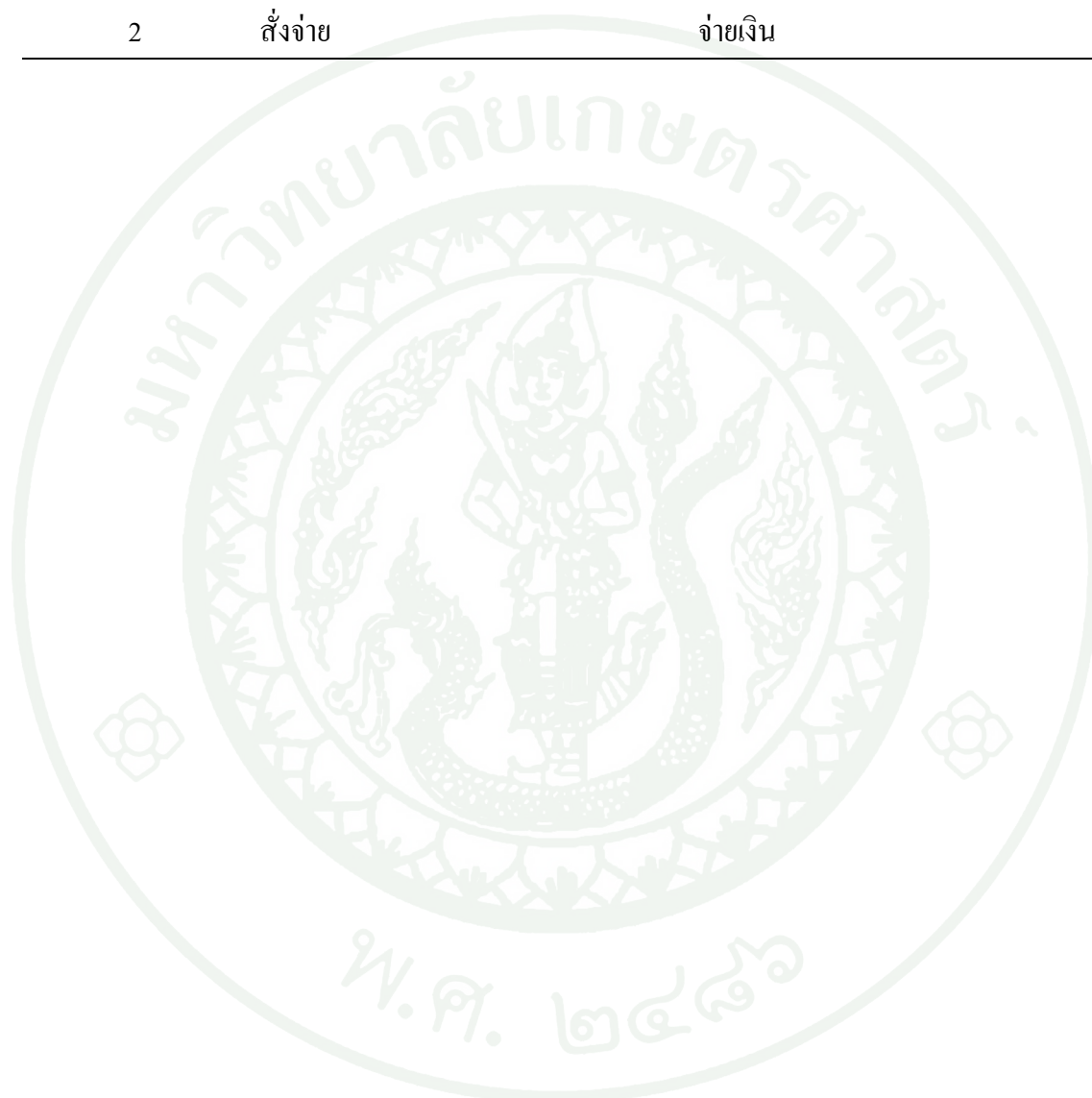
ลำดับ	ออนโทโลยีระดับ 1	ออนโทโลยีระดับ 2
5	เพิกถอน	การสละมรดก โอน จีดม่า ทำลาย สละพินัยกรรม
6	หลายคน	เสียงข้างมาก
7	ชำระหนี้	ปลดหนี้ โอนสิทธิ
8	แบ่งมรดก	แบ่งปัน ตัดส่วน สัญญาประนีประนอม
9	จัดการ	โดยลำพัง เสียงข้างมาก สืบหา
10	ไม่สมบูรณ์	บุคคล ตก เดิม แก้ไข

ตารางผนวกที่ ก4 คำในโนคของออนโทโลยีตัวเงินที่ใช้ในขั้นตอนวิธี SCRO II

ลำดับ	ออนโทโลยีระดับ 1	ออนโทโลยีระดับ 2
1	ผู้ทรง	ผู้ทรงเช็ค ผู้ถือ ผู้รับเงิน
2	เจ้าหนี้	หนี้

ตารางผนวกที่ ก5 คำในความสัมพันธ์ของออนโทโลยีตัวเงินที่ใช้ในขั้นตอนวิธี SCRO II

ลำดับ	ออนโทโลยีระดับ 1	ออนโทโลยีระดับ 2
1	อาวัล	ประกัน
2	สั่งจ่าย	จ่ายเงิน





การประชุมวิชาการ และแสดงผลงานทางวิชาการ พระจอมเกล้าลาดกระบัง ประจำปี 2552

สำนักบริหารวิชาการ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

31 สิงหาคม – 2 กันยายน 2552

การใช้ออนโทโลยีเพื่อการสืบค้นคำพิพากษาศาลฎีกา Supreme Court sentences Retrieval using Ontology

ธนพล ดันติศรีปรีชา¹, นววรรณ สุนทรนิษฐ์²

¹นักศึกษา ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ E-mail: g5164138@ku.ac.th

²ผู้ช่วยศาสตราจารย์ ประจำภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ E-mail:

fscinws@ku.ac.th

บทคัดย่อ

การสืบค้นคำพิพากษาศาลฎีกาโดยใช้ออนโทโลยี (ontology) มีวัตถุประสงค์เพื่อสืบค้นคำพิพากษาศาลฎีกาให้ได้คำพิพากษาศาลฎีกาที่มีความหลากหลาย ผู้วิจัยได้นำเสนอวิธี SCRO โดยการใช้ออนโทโลยีมาเพิ่มกลุ่มของคำค้นที่รับมาจากผู้ใช้ และทำการกำหนดค่าสัมประสิทธิ์ของแต่ละคำค้นเพื่อให้คำค้นมีความสำคัญที่แตกต่างกัน จากผลการทดลองแสดงให้เห็นว่าวิธีการ SCRO สามารถเพิ่มประสิทธิภาพในการสืบค้นคำพิพากษาศาลฎีกา ให้ได้คำพิพากษาศาลฎีกาที่มีความหลากหลาย ได้ดีกว่าวิธีการสืบค้นในปัจจุบัน

คำสำคัญ: ออนโทโลยี, ขยายคำค้น, สืบค้นสารสนเทศ, คำพิพากษาศาลฎีกา

Abstract

The Objective of Supreme Court Sentences Retrieval algorithm is to provide a set of diversify documents for user. We proposed a new algorithm called SCRO which use ontology in order to do query expansion. The set of keywords generated from our algorithm are assigned with weights calculated by the system. The experimental result shows that SCRO algorithm can enhance the performance of the retrieval system. The outputs of the retrieved system are evaluated by law expert and we found that all documents are satisfied by the user.

Keywords: ontology, query expansion, information retrieval, Supreme Court sentences.

1. บทนำ

การสืบค้นเอกสารที่อยู่ในกลุ่ม (domain) เดียวกันในปัจจุบัน นิยมใช้คำค้น (keyword) [1] จากผู้ใช้เพียงอย่างเดียว เนื่องจากสามารถทำได้ง่าย และรวดเร็วในการเขียนโปรแกรม แต่ข้อเสียของวิธีการนี้คือไม่สามารถจัดลำดับเอกสารที่สืบค้นได้อย่างถูกต้อง เนื่องจากเอกสารที่อยู่ในกลุ่มเดียวกัน มีลักษณะของคำที่ปรากฏอยู่คล้ายกัน ทำให้การสืบค้นแบบเดิมโดยใช้ความถี่ของคำเป็นตัวจัดลำดับ ไม่สามารถจัดลำดับ

เอกสารและสืบค้นเอกสารออกมาได้ตรงกับความต้องการของผู้ใช้

ระบบการสืบค้นคำพิพากษาในปัจจุบัน [2] รับคำค้นจากผู้ใช้ในการสืบค้นเพียงอย่างเดียว ทำให้คำพิพากษาที่ได้อาจไม่ตรงกับความต้องการของผู้ใช้ เนื่องจากผู้ใช้ที่ใช้ระบบสืบค้นคำพิพากษาในปัจจุบัน บางกลุ่ม เป็นบุคคลทั่วไปที่ไม่มีความรู้ทางด้านกฎหมาย และต้องการค้นคืนคำพิพากษาศาลฎีกาเพื่อให้อ้างอิงถึงลักษณะการกระทำผิดทางกฎหมายในประเด็นต่างๆ ซึ่งคำพิพากษาศาลฎีกาที่สืบค้นได้ไม่มีความหลากหลายในประเด็นทางกฎหมาย ทำให้ผู้ใช้

ความสัมพันธ์ที่เท่ากัน ทำให้ผลลัพธ์ของการสืบค้นคำพิพาทศาสตร์ได้คำพิพาทศาสตร์ที่เนื้อหาไม่มีความหลากหลาย

การสืบค้นคำพิพาทศาสตร์ในเรื่องที่อยู่ในกลุ่มเดียวกัน ลักษณะของคำพิพาทศาสตร์จะมีความหมายใกล้เคียงกัน ทำให้การสืบค้นคำพิพาทศาสตร์โดยการเปรียบเทียบจากคำค้นเพียงคำเดียว ได้คำพิพาทศาสตร์ที่ไม่มี ความหลากหลายในข้อเท็จจริงของคดี

2.2 ออนโทโลยี

ออนโทโลยี คือ โครงสร้างข้อมูลชนิดกราฟที่ใช้เพื่อแทนความรู้ในรูปแบบโครงข่ายความหมาย ซึ่งประกอบด้วยโนด (node) และเส้นเชื่อม โดยโนดใช้ในการอธิบายความหมาย (concept) และเส้นเชื่อมใช้ในการอธิบายความสัมพันธ์ (relationship) ระหว่างคำในออนโทโลยี [3] โดยตัวอย่างโครงสร้างออนโทโลยีแสดงดังรูปที่ 1

ประโยชน์ของเอกสารแต่ละย่อหน้า ประกอบไปด้วยคำหลายๆ คำที่มีความสัมพันธ์กัน การสืบค้นโดยใช้เพียงคำค้นจากผู้ใช้ คำที่รับมานั้นอาจไม่ปรากฏในเอกสารที่ตรงกับความต้องการ ดังนั้น การใช้ออนโทโลยีมาช่วยในการสืบค้นเอกสารในขั้นตอนการการเพิ่มจำนวนคำค้นเพื่อการสืบค้นเอกสาร (query expansion) ทำให้สามารถครอบคลุมแต่ละคำที่เกิดขึ้นในแต่ละประโยค แต่วิธีการนี้ขึ้นอยู่กับจำนวนของคำที่เพิ่มขึ้นมา และชนิดของคำที่ปรากฏบนเอกสาร

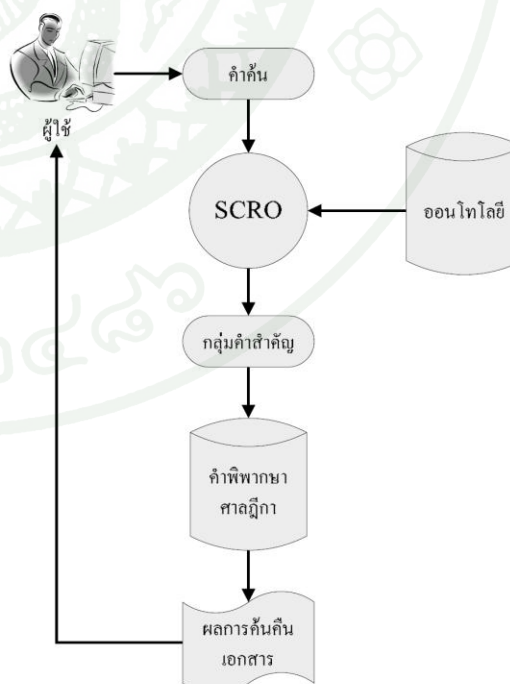
งานวิจัยของ Oifa Dridi [4] ได้นำออนโทโลยีมาประยุกต์ใช้ในการสืบค้นข้อมูลโรคกระเพาะเรื้อรัง เนื่องจากการใส่คำค้นโดยตรงจากผู้ใช้ ไม่สื่อความหมายที่สามารถจะสืบค้นในเอกสารที่ต้องการได้ ซึ่งระบบนี้ได้นำออนโทโลยีโรคกระเพาะเรื้อรังมาใช้ในการดูความสัมพันธ์ของคำ โดยผลการทดลองสามารถค้นคืนเนื้อหาที่ตรงกับความต้องการของผู้ใช้ได้

ในการคำนวณค่าน้ำหนักของคำแต่ละคำในออนโทโลยี งานวิจัยหลายๆ งานได้ใช้ความถี่เป็นตัว

กำหนดค่าน้ำหนักของแต่ละคำ โดยหาจากสูตร $TF*IDF$ โดยจะมีการหาค่าความถี่ที่แตกต่างกันขึ้นอยู่กับชนิดของข้อมูล Jibran Mustafa [5] ได้เสนอการหาค่าความถี่ โดยคำนวณจากความถี่ของคำที่ปรากฏในเอกสารด้วยความถี่สูงสุดของคำๆ นั้นที่ปรากฏในแต่ละเอกสาร

3. ขั้นตอนวิธี SCRO

งานวิจัยนี้ได้ทำการวิจัยเกี่ยวกับ การสืบค้นคำพิพาทศาสตร์ จากประมวลกฎหมายแพ่งและพาณิชย์ ในส่วนของ บรรพ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน โดยนำออนโทโลยีที่ได้จากผู้เชี่ยวชาญมาใช้เป็นฐานความรู้เพื่อเพิ่มคำค้นที่ได้รับจากผู้ให้ และนำคำที่ได้มาสืบค้นคำพิพาทศาสตร์ที่ได้เตรียมไว้ ซึ่งมีแผนภาพการทำงานดังรูปที่ 2 โดยวิธีเพิ่มสัมพันธ์ของความสำคัญระหว่างคำสำคัญในออนโทโลยีเปรียบเทียบกับ การสืบค้นโดยวิธีเดิม



ภาพที่ 2: ขั้นตอนการสืบค้น

วิธีการสร้างออนโทโลยีคำพินาศภาษาศาสตร์ ผู้เชี่ยวชาญจะเป็นผู้กำหนดโนดที่ใช้เป็นตัวแทนของคำ และเชื่อมความสัมพันธ์ระหว่างโนด โดยโนดเหล่านี้เป็นความสัมพันธ์ที่เกิดขึ้นจากประมวลกฎหมายแพ่งและพาณิชย์ และถูกสกัดออกมาเพื่อใช้ในการสร้างออนโทโลยี

การเพิ่มจำนวนคำค้นเพื่อการสืบค้นเอกสาร (query expansion) เป็นกระบวนการที่ระบบจะหาคำที่มีความสัมพันธ์กับคำค้นที่ผู้ใช้ได้ป้อนเข้ามาเพื่อใช้ในการสืบค้น โดยคำที่เพิ่มเข้ามาจะเป็นคำที่มีความหมายเหมือนกัน หรือเป็นคำที่มีความหมายตรงกันข้าม หลังจากนั้นก็จะนำชุดของคำค้นที่เข้าไปสืบค้นเอกสารที่มีชุดของคำแต่ละคำปรากฏอยู่ ตัวอย่างผู้ใช้ป้อนคำว่า “ผู้ทรง” ระบบจะทำการเพิ่มคำโดยอัตโนมัติได้เป็น “ผู้รับเงิน” “ผู้ถือ” “ผู้รับเงิน” “ผู้ถือ” “โล่เบี้ย” “วันออกเช็ก” “สลักหลัง” “ตัวเงิน” “เช็ก” “ตัวแลกเงิน” “ตัวสัญญาใช้เงิน”

ในงานวิจัยนี้ ใช้ออนโทโลยีในการเพิ่มคำสำคัญ โดยนำคำค้นที่ได้จากผู้ใช้เข้าไปค้นหาในออนโทโลยี และนำคำที่มีความสัมพันธ์กับคำค้นที่เพิ่มเข้ามาในชุดของคำค้น โดยจะค้นลงไปภายในออนโทโลยี 1 ระดับ และเมื่อคำที่ถูกค้นขึ้นมีความสัมพันธ์เชื่อม “คือ” ก็ จะทำการค้นหาลงไปอีก 1 ระดับ จนได้ครบทุกคำที่ต้องการ

3.1 การให้น้ำหนักของคำ

ค่าน้ำหนักของคำค้นในแต่ละคำพินาศภาษาศาสตร์ คือ ค่าที่ใช้บอกความใกล้เคียงกันของคำค้น [6, 7] ในแต่ละคำพินาศภาษาศาสตร์กับคำพินาศภาษาศาสตร์ที่ต้องการสืบค้น โดยค่าของน้ำหนักของคำค้นหาจากสูตรที่ (1)

$$w_{i,j} = tf_{i,j} * idf_j \quad (1)$$

โดยที่ $w_{i,j}$ คือ ค่าน้ำหนักของคำ i ที่ปรากฏในคำพินาศ j

$tf_{i,j}$ คือ ความถี่ของคำ i ที่ปรากฏในคำพินาศ j

idf_j คือ ค่าที่ใช้วัดการกระจายของคำในแต่ละเอกสาร j

จากสูตรที่ (1) เราหาค่า $tf_{i,j}$ ได้จากสูตรที่ (2)

$$tf_{i,j} = \frac{f_{i,j}}{\max(f_i)}$$

โดยที่ $f_{i,j}$ คือ ความถี่ของคำ i ที่ปรากฏในคำพินาศ j

$\max(f_i)$ คือ ค่าความถี่สูงสุดของคำ i

และค่า idf_j จากสูตรที่ (1) หาได้จาก

$$idf_j = \ln\left(\frac{N}{n_j}\right)$$

โดยที่ N คือ จำนวนเอกสารทั้งหมด

n_j คือ จำนวนเอกสารที่มีคำ i ปรากฏอยู่

3.2 การเรียงลำดับเอกสาร

หลังจากได้ค่าน้ำหนักของคำค้นในแต่ละคำพินาศภาษาศาสตร์แล้ว จะทำการเรียงลำดับของคำพินาศภาษาศาสตร์ โดยนำค่าน้ำหนักของแต่ละคำค้นคูณค่าสัมประสิทธิ์ความสำคัญของคำค้นที่ได้จากออนโทโลยี และรวมค่าน้ำหนักของคำสำคัญที่คูณด้วยสัมประสิทธิ์ความสำคัญทั้งหมดไว้ด้วยกัน ซึ่งผลลัพธ์ที่ได้ คือ ค่าน้ำหนักรวมของคำ ที่ใช้วัดลำดับการแสดงผลของเอกสารให้กับผู้ใช้

ตารางที่ 1: กฎในการกำหนดสัมประสิทธิ์ความสำคัญ

ความสัมพันธ์	สัมประสิทธิ์ความสำคัญของโนด	
	โนดถูก	โนดแม่
คือ	ค่าสัมประสิทธิ์ลดลง $\frac{1}{4}$ ของค่าสัมประสิทธิ์ปัจจุบัน	ค่าสัมประสิทธิ์ลดลง $\frac{1}{4}$ ของค่าสัมประสิทธิ์ปัจจุบัน
อื่นๆ	ค่าสัมประสิทธิ์ลดลง $\frac{1}{2}$ ของค่าสัมประสิทธิ์ปัจจุบัน	ค่าสัมประสิทธิ์ลดลง $\frac{1}{4}$ ของค่าสัมประสิทธิ์ปัจจุบัน

ความสัมพันธ์ระหว่างคำค้นในคำพินาศภาษาศาสตร์จะเป็นตัวบอกค่าสัมประสิทธิ์ระหว่างคำค้น 2 คำค้น โดยผู้เชี่ยวชาญจะเป็นผู้กำหนดค่าสัมประสิทธิ์ โดยดูจากความสัมพันธ์ของคำในออนโทโลยี โดยที่ค่าสัมประสิทธิ์เท่ากับ 1 เมื่อคำนั้นเป็นคำค้นหลัก และ

การประชุมวิชาการ และแสดงผลงานทางวิชาการ พระจอมเกล้าลาดกระบัง ประจำปี 2552

สำนักบริหารวิชาการ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

31 สิงหาคม – 2 กันยายน 2552

แบ่ง 2 กรณีในการหาค่าสัมประสิทธิ์ ตามกฎในการกำหนดสัมประสิทธิ์ความสำคัญ แสดงในตารางที่ 1 เมื่อได้ค่าสัมประสิทธิ์ความสำคัญ จะนำมาคำนวณหา น้ำหนักของค่าสำคัญแต่ละค่าใหม่ตามสูตรที่ (4)

$$Weight_{i,j} = \sum w_{i,j} * k_i \quad (4)$$

ค่าน้ำหนักที่ได้จากสูตรที่ (4) จะเป็นค่าน้ำหนักรวมของค่าที่เพิ่มขึ้นมาจากออนโทโลยี คุณด้วยค่าสัมประสิทธิ์ความสำคัญ ซึ่งค่าสัมประสิทธิ์จะเป็นตัวกำหนดความสำคัญของค่าแต่ละค่าที่ได้จากออนโทโลยี ทำให้ในการสืบค้น ค่าทุกค่าจะมีความสำคัญไม่เท่ากัน โดยค่าที่มีความหมายใกล้เคียงกัน หรือใกล้เคียงกับค่าอื่น จะมีค่าความสำคัญมากกว่า ค่าที่มีความหมายต่างกับค่าอื่น

4. ขั้นตอนและวิธีการทดลอง

4.1 ข้อมูลทดสอบ

ข้อมูลที่ใช้ในการทดลองสืบค้นคำพิพาทภาษาศาลฎีกาในงานวิจัยนี้ได้รวบรวมมาจาก คำพิพาทภาษาศาลฎีกาจากเว็บไซต์ ระบบสืบค้นคำพิพาทและคำสั่งคำร้องศาลฎีกา [1] โดยรวบรวมคำพิพาทภาษาศาลฎีกาไว้ทั้งหมด 200 คำพิพาทของศาลฎีกาที่แตกต่างกันโดยเกี่ยวกับบรรพ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน ตั้งแต่มาตราที่ 898 ถึงมาตราที่ 1011 จากคำพิพาทศาลฎีกาทั้งหมด 99,656 คำพิพาท ซึ่งคำพิพาทเหล่านี้จะมีเนื้อหาที่มีลักษณะใกล้เคียงกันในแต่ละคำพิพาท คำค้นที่ใช้ทั้งหมด 10 คำ ดังตารางที่ 2 ซึ่งแต่ละคำค้นเป็นคำที่มีอยู่ในออนโทโลยีตัวเงิน

ตารางที่ 2: คำค้นที่ใช้ในการทดลอง

คำค้นหลักจากผู้ใช้	คำค้นที่เพิ่มขึ้นจากขั้นตอน SCRO
ผู้ทรง	ผู้รับเงิน ผู้ถือ ไลเบียร์ วันออกเช็ค สลากหลัง ตัวเงิน เช็ค ตัวแลกเงิน ตัวสัญญาใช้เงิน
ตัวเงิน	เช็ค ตัวแลกเงิน ตัวสัญญาใช้เงิน โอน ผู้ทรง ผู้รับเงิน ผู้ถือ
โอน	ส่งมอบ สลากหลัง ตัวเงิน เช็ค ตัวแลกเงิน ตัวสัญญาใช้เงิน

เช็ค	ตัวเงิน ธนาคาร ชัดคร่อมทั่วไป ชัดคร่อมเฉพาะ ขาดลด แกล้งปลอม ลายมือ มิได้ประทับตรา
ตัวแลกเงิน	ตัวเงิน ผู้ส่งจ่าย ผู้จ่าย
ตัวสัญญาใช้เงิน	ตัวเงิน ใช้เงินเมื่อทางถาม
แกล้ง	ปลอมลายมือ มิได้ประทับตรา เช็ค ตัวเงิน ตัวแลกเงิน ตัวสัญญาใช้เงิน
ผู้ถือ	ผู้ทรง ไลเบียร์ วันออกเช็ค สลากหลัง
สลากหลัง	ผู้ทรง ผู้ถือ ผู้รับเงิน โอน ส่งมอบ อวาล์
ไลเบียร์	ผู้ทรง ผู้ถือ ผู้รับเงิน

จากตารางที่ 2 ช่องคำค้น คือ คำค้นที่ใช้ในการทดลอง ซึ่งใช้ทั้งหมดจำนวน 10 คำค้น และช่องคำค้นที่เพิ่มขึ้น คือ คำค้นที่เพิ่มขึ้นจากความสัมพันธ์ของออนโทโลยีตัวเงินที่ผู้เชี่ยวชาญได้กำหนดไว้

4.2 ขั้นตอนการทดลอง

การสืบค้นคำพิพาทภาษาศาลฎีกา โดยใช้ออนโทโลยีในการเพิ่มคำสำคัญจากความสัมพันธ์ในออนโทโลยีตัวเงิน โดยมีขั้นตอนการทดลอง ดังนี้

1) ผู้เชี่ยวชาญสร้างออนโทโลยีตัวเงิน โดยขั้นตอนนี้ผู้เชี่ยวชาญจะสกัดคำสำคัญจากคำพิพาทภาษาศาลฎีกา และประมวลกฎหมายแพ่งและพาณิชย์ ในส่วนบรรพ 3 เอกเทศสัญญา ลักษณะ 21 ตัวเงิน มาสร้างเป็น โครงข่ายของความสัมพันธ์กัน โดยผู้เชี่ยวชาญจะกำหนดความสัมพันธ์ และคำที่ปรากฏในออนโทโลยีตัวเงิน

2) ขั้นตอนในการเพิ่มคำค้นของคำพิพาทศาลฎีกา ขั้นตอนนี้จะนำคำค้นที่เตรียมไว้เข้าไปค้นหาในออนโทโลยีตัวเงิน และดึงคำค้นที่มีความสัมพันธ์กับคำค้นหลักนั้นมาเพิ่มให้ชุดของคำค้น โดยทำการค้นหาจนกว่าไม่มีความสัมพันธ์ที่เชื่อมด้วย “คือ” ผลลัพธ์ที่ได้คือ ชุดของคำค้นใหม่ เพื่อใช้ในการสืบค้นคำพิพาทภาษาศาลฎีกาจากฐานข้อมูลคำพิพาทภาษาศาลฎีกาที่เตรียมไว้

3) เรียงลำดับคำพิพาทภาษาศาลฎีกาที่สืบค้นได้ โดยขั้นตอนนี้จะคำนวณค่าน้ำหนักของแต่ละ

การประชุมวิชาการ และแสดงผลงานทางวิชาการ พระจอมเกล้าลาดกระบัง ประจำปี 2552

สำนักบริหารวิชาการ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

31 สิงหาคม – 2 กันยายน 2552

คำค้นที่ปรากฏในแต่ละคำพิพาทศาสตร์ศึกษา และนำมาคูณกับค่าสัมประสิทธิ์ที่ผู้เชี่ยวชาญได้กำหนดไว้ในขั้นตอนการสร้างออนโทโลยีตัวเงิน

5. ผลการทดลอง

ผู้วิจัยได้ทำการทดลองเพื่อเปรียบเทียบวิธีการสืบค้นแบบเดิม คือ ใช้คำค้นเพียง 1 คำจากผู้ใช้และวิธี SCRO คือ เพิ่มชุดของคำค้นจากคำค้นหลักโดยใช้ออนโทโลยีตัวเงินเพิ่มคำค้น และผู้เชี่ยวชาญทำการประเมินผลลัพธ์ของการสืบค้นคำพิพาทศาสตร์ศึกษาโดยพิจารณาจาก

- 1) คำพิพาทศาสตร์ศึกษาที่สืบค้นได้มีประเด็นทางกฎหมายที่ต้องวินิจฉัยสอดคล้องกับคำค้นหลักหรือไม่
- 2) การสืบค้นผลลัพธ์ทั้ง 5 อันดับ ต้องมีมุมมองข้อเท็จจริงที่แตกต่างกัน แต่ยังคงมีประเด็นหลักตรงตามข้อ 1

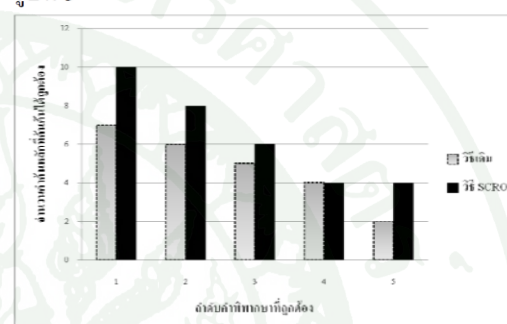
ผลการทดลองจะทำการพิจารณาจำนวนของคำพิพาทศาสตร์ศึกษา ตั้งแต่ 1 อันดับแรก ไปจนถึง 5 อันดับแรกที่มีความแตกต่างกันในด้านของเนื้อหาความในคำพิพาทศาสตร์ศึกษา ซึ่งผู้วิจัยได้ทำการทดลอง และให้ผู้เชี่ยวชาญทำการประเมินผลได้ผลลัพธ์ดังตารางที่ 3

ตารางที่ 3: ผลการประเมินจากผู้เชี่ยวชาญตาราง

ลำดับเอกสารที่ ถูกต้อง	ผลการประเมินจากผู้เชี่ยวชาญ (%)	
	วิธีเดิม	วิธี SCRO
1 อันดับแรก	70	100
2 อันดับแรก	60	80
3 อันดับแรก	50	60
4 อันดับแรก	40	40
5 อันดับแรก	20	40

จากตารางที่ 3 วิธีเดิมให้ผลลัพธ์การค้นคือคำพิพาทศาสตร์ศึกษาที่สื่อความหมายตรงกับคำค้นใน 1 อันดับแรก 70 เปอร์เซ็นต์ ส่วนวิธี SCRO วิธีเดิมให้ผลลัพธ์การค้นคือคำพิพาทศาสตร์ศึกษาที่สื่อความหมายตรงกับคำค้นใน 1 อันดับแรก 100 เปอร์เซ็นต์ ซึ่งวิธี

SCRO ให้ผลลัพธ์ของการสืบค้นคำพิพาทศาสตร์ศึกษาที่ถูกต้องได้ดีกว่าวิธีเดิม และเป็นเช่นนี้ ไปจนถึง 5 อันดับแรกของการสืบค้นคำพิพาทศาสตร์ศึกษาเรื่องตัวเงิน โดยสามารถแสดงประสิทธิภาพของการสืบค้นคำพิพาทศาสตร์ศึกษากระหว่างวิธีเดิมและ วิธี SCRO ได้ดังรูปที่ 3



รูปที่ 3: กราฟเปรียบเทียบประสิทธิภาพระหว่างวิธีเดิมและวิธี SCRO

จากรูปที่ 3 เมื่อพิจารณาจำนวนคำค้นที่สืบค้นคำพิพาทศาสตร์ศึกษาได้ถูกต้องโดยพิจารณาจากเกณฑ์ที่ผู้เชี่ยวชาญได้ตั้งไว้ ใน 3 อันดับแรก วิธีเดิม สามารถสืบค้นได้ถูกต้อง 50 เปอร์เซ็นต์ ในขณะที่ วิธี SCRO สามารถสืบค้นได้ถูกต้อง 60 เปอร์เซ็นต์ และจำนวนของคำค้นที่สืบค้นได้ถูกต้องมีแนวโน้มลดลงทั้ง 2 วิธีเมื่อลำดับเอกสารที่ใช้ในการเปรียบเทียบเพิ่มมากขึ้น และแสดงให้เห็นว่าวิธี SCRO สามารถสืบค้นคำพิพาทศาสตร์ศึกษาที่มีความหลากหลายได้มากกว่าวิธีเดิมในทุกๆ 5 เอกสารแรก

6. สรุป

งานวิจัยนี้ ได้นำเสนอขั้นตอนวิธี SCRO เพื่อใช้ในการสืบค้นคำพิพาทศาสตร์ศึกษา เพื่อให้ได้คำพิพาทศาสตร์ที่ตรงกับความต้องการ และผลลัพธ์ของการสืบค้นคำพิพาทศาสตร์มีความหลากหลาย จากผลการทดลองแสดงให้เห็นว่า การสืบค้นโดยทำการเพิ่มคำค้นจากออนโทโลยี สามารถสืบค้นคำพิพาทศาสตร์ศึกษาที่เนื้อหามีความหลากหลาย มากกว่าวิธีการค้นโดยใช้เพียงคำค้นของผู้ใช้เพียงอย่างเดียว โดยสามารถ

การประชุมวิชาการ และแสดงผลงานทางวิชาการ พระจอมเกล้าลาดกระบัง ประจำปี 2552

สำนักบริหารวิชาการ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

31 สิงหาคม – 2 กันยายน 2552

สืบค้นคำพิพากษาได้ประสิทธิภาพในทุกๆ 5 อันดับแรกของเอกสารที่ตรงกับความต้องการ เนื่องจากวิธีนี้ทำการเพิ่มคำค้นโดยอาศัยโครงสร้างความสัมพันธ์จากออนโทโลยี มาใช้ในการเพิ่มคำค้น และใช้ออนโทโลยีในการกำหนดสัมประสิทธิ์ความสำคัญ ทำให้คำค้นที่ได้เพิ่มมานั้น มีความสำคัญแตกต่างกันไปตามระดับของความสัมพันธ์ในออนโทโลยีตัวเงิน

the IEEE EMBS Cité Internationale, Lyon, France, August 23-26, 2007.

- [7] J. Fiaidhi, S. Mohammed, J. Jaam and A. Hasnah, "A Standard Framework for Personalization Via Ontology-Based Query Expansion", 2nd Pakistan Journal of Information and Technology, Pakistan, 2003

เอกสารอ้างอิง

- [1] M. Kobayashi and K. Takeda, "Information Retrieval on the Web", ACM Computing Surveys (CSUR), June, 2000.
- [2] เว็บไซต์ระบบสืบค้นคำพิพากษาและคำสั่งคำร้อง ศาลฎีกา <http://www.deka2007.supremecourt.or.th/>
- [3] Agissilaos Andreou, "Ontologies and Query expansion", 2005.
- [4] O. Dridi and M. Ben Ahmed, "Building an Ontology-Based Framework For Semantic Information Retrieval: Application To Breast Cancer", 3rd International Conference, April 7-11, 2008.
- [5] J. Mustafa, S. Khan and K. Latif, "Ontology Based Semantic Information Retrieval", 4th International IEEE Conference "Intelligent Systems", 2008
- [6] C. Bratsas, V. Koutkias, E. Kaimakamis, P. Bamidis, and N. Maglaveras, "Ontology-based Vector Space Model and Fuzzy Query Expansion to Retrieve Knowledge on Medical Computational Problem Solutions" Proceedings of the 29th Annual International Conference of

ขั้นตอนวิธีการขยายคำค้นสำหรับการค้นคืนคำพิพากษาศาลฎีกาด้วยออนโทโลยี

Query Expansion Algorithm for Supreme Court Sentences Retrieval using Ontology

ธนพล ตันติศรีปรีชา¹ และนวลวรรณ สุนทรภิชัย¹

Tanapon Tantisripreecha¹ and Nuanwan Soonthornphisaj¹

บทคัดย่อ

การสืบค้นคำพิพากษาศาลฎีกาโดยใช้ออนโทโลยี (ontology) มีวัตถุประสงค์เพื่อสืบค้นคำพิพากษาศาลฎีกาเพื่อให้ได้คำพิพากษาศาลฎีกาที่ประเด็นทางกฎหมายมีความหลากหลาย ผู้วิจัยได้นำเสนอขั้นตอนวิธี SCRO โดยการใช้ออนโทโลยีในการเพิ่มกลุ่มของคำค้นที่รับมาจากผู้ใช้ และทำการกำหนดค่าสัมประสิทธิ์ของแต่ละคำค้นเพื่อให้คำค้นมีความสำคัญที่ต่างกัน ผู้วิจัยได้ทำการทดลองกับชุดข้อมูลจากคำพิพากษาศาลฎีกา 2 ชุด คือ ชุดข้อมูลมรดก และ ชุดข้อมูลตัวเงิน ผลการทดลองแสดงให้เห็นว่าขั้นตอนวิธี SCRO สามารถเพิ่มประสิทธิภาพในการสืบค้นคำพิพากษาศาลฎีกา โดยพบว่าคำพิพากษาศาลฎีกามีประเด็นทางกฎหมายที่มีความหลากหลาย ได้ดีกว่าวิธีการสืบค้นในระบบการค้นคืนคำพิพากษาในปัจจุบัน

ABSTRACT

This study presents an ontological approach for Supreme Court Sentence Retrieval. A legal ontology is applied in order to do query expansion in order to retrieve a set of diversified law issues for user. A set of queries are generated by the algorithm, called SCRO, and applied for the retrieval process. We perform the experiments on two datasets which are Supreme Court Sentence related to succession law and bill of exchange law. We compare the performance of our algorithm to the keyword-based approach and found that the proposed algorithm outperforms the traditional searching algorithm in term of document diversity.

Keywords: Ontology, Query expansion, Legal document retrieval.

T Tantisripreecha: g5164138@ku.ac.th

Introduction

Supreme Court sentences are useful for law students, lawyers and judges. These people study Supreme Court sentences in order to learn new precedent that cannot be found in the code law. A domain specific search engine was implemented as a tool to retrieve all related Supreme Court sentences based on keywords search methodology.

The problem found in the current system is that a set of retrieved documents lack of diversity therefore the objective of this research is to develop a new algorithm using ontology for the query expansion process in order to retrieve a set of documents with diversified law issues. These diversified law issues can be used to fulfill knowledge for those law users.

Another advantage of our research is that user does not need to input several keywords for searching because our algorithm has an ability to generate a set of queries and produce the most suitable Supreme Court sentences to user.

Ontology

An ontology is a semantic network which consists of a set of nodes and links. It is used for knowledge representation in many works. Many ontologies were constructed such as breast cancer ontology [1]. Wu, F., *et al.* [2] proposed an algorithm using ontology for query expansion. They classified a set of links into 2 classes called related links and example links. Benjamin, *et al.* [3] implemented a legal ontology to improve IT Support for professional judges in Spain. They used questionnaires as a tool to collect requirements from young Spanish judges in year 2002. After the survey, all questionnaires were analysed and were implemented as an ontology. Ontology was applied in e-learning domain by Lee, *et al.* [4]. They used ontology as a tool for learning object retrieval. The digital learning materials were retrieved and evaluated in term of relevant. The experimental results showed that the ontology can be used to enhance the performance. Saravanan, *et al.* [5] did research about legal information retrieval using ontology. They applied ontology for the query expansion process to enhance the retrieval performance. The experiment was set up on Indian laws related to rent control, income tax and sale tax. They compared the proposed algorithm to the keyword search and found that their algorithm outperformed the traditional searching algorithm. Guiraud [6] proposed an ontology updating algorithm for French law. They used natural language processing technique in the updating process. The objective of their work was to keep updating ontology according to the law changing.

The Supreme Court sentences retrieval process is started for each query, q_i . Considering a set of retrieved documents, the SCRO algorithm ranks these documents using TFIDF formular (see equation 1)

$$w_{in} = tf_{in} \times idf_i \quad (1)$$

where w_{in} = weight value of word _{i} in the retrieved document _{n}
 tf_{in} = the number of occurrences for word _{i} in document _{n}

$$idf = \ln\left(\frac{N}{n_i}\right) \quad (2)$$

where N = the number of retrieved documents
 n_i = the number of retrieved documents contain word _{i}

The SCRO algorithm selects the first ranked document as an output for each query and checks for duplicated documents. Since, each query is processed independently, therefore the duplicate checking process is done to finalize the output. Because the objective of the SCRO algorithm is to ensure the diversity of law issues from the final set of retrieved documents.

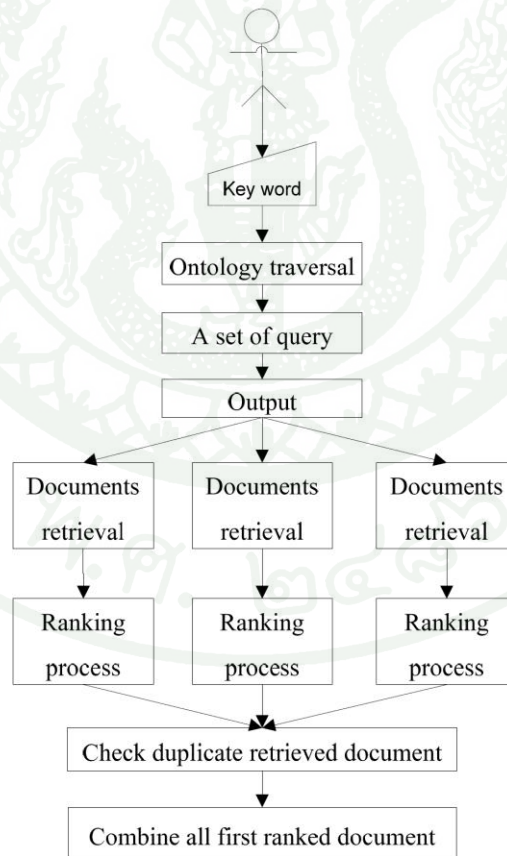


Fig. 3. The framework of SCRO algorithm

Algorithm : SCRO

Input: Ontology, keyword

Begin

Q = OntologyTraverse(Ontology, Keyword)

For each q_i in QD_i = DocumentRetrieve(q_i)Ranking(D_i)

End For

CombineResults = checkDuplicate()

End

Output: CombineResults

Fig. 4. SCRO algorithm

Baseline algorithm

We compare the performance of SCRO algorithm with the current system implemented by the Thai Supreme Court. The system can be accessed via the website www.deka2007.supremecourt.or.th

The current system uses term frequency as a mechanism for the keyword search. The set of retrieved documents obtained from the baseline algorithm will be evaluated base on the diversity of law issues.

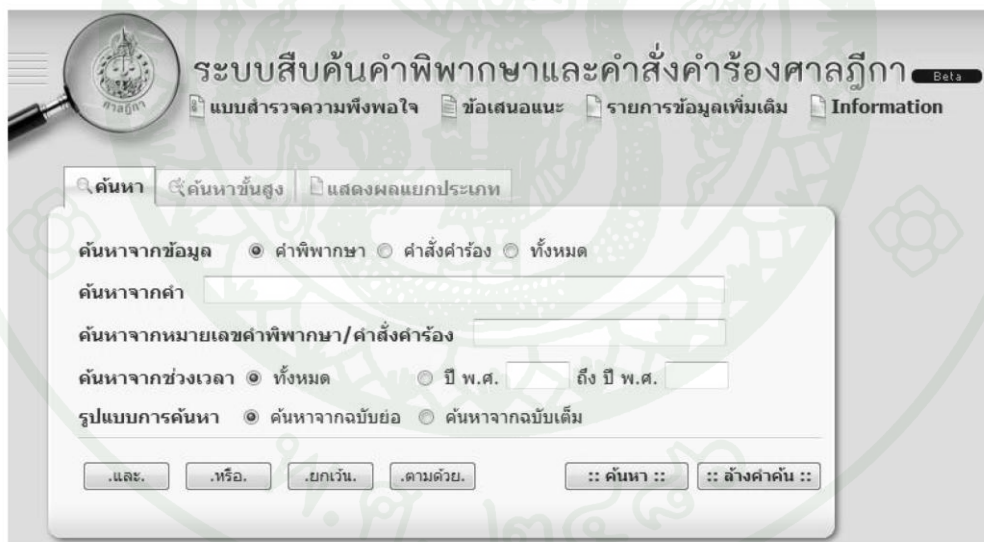


Fig. 5. The Supreme Court Sentences Retrieval System

Performance Measurement

A set of retrieved documents are measured based on the degree of diversity coverage (see equation 3). The maximum value of diversity is 1 that means the algorithm can retrieve all different law issues related to the keyword.

$$DC_{(w)} = \frac{\text{number of different law issues obtained from algorithm}}{\text{total number of different law issues found in ontology}} \quad (3)$$

The data set used in our experiment is the Supreme Court Sentences related to the Thai succession law and bill of exchange.

For each searching keyword, we get the number of different law issues from the predefined ontology using SCRO algorithm. Next, the retrieval module is performed to get the results. For example, the keyword, “มรดก”, is related to 13 law issues. Therefore the retrieval module retrieves 13 court sentences for user. We expect that these 13 documents should have different law issues. By the way, they might have some documents which have the same law issue. Finally, we calculate the precision in term of diversity value using equation 3.

Experimental Results

1 The first experiment

In order to evaluate the retrieval performance on Thai succession law, we set up the experiment using 10 keywords which are มรดก (estate), ผู้จัดการมรดก (administrator of an estate), ทายาท (heir), ศาล (court), ผู้ตาย (deceased), พยาน (witness), พินัยกรรม (wills), บุตรนอกกฎหมาย (illegitimate child), ถอน (revoke), เจ้าหนี้ (creditor). The dataset consists of 493 Court Sentences related to the succession law. The performance compared between SCRO algorithm and the baseline algorithm are shown in Table 1.

Table 1. The performance comparison between SCRO algorithm and the baseline algorithm

keyword	NO. of law issues		Total law issues	Diversity (DC)	
	SCRO	Baseline		SCRO	Baseline
มรดก (estate)	11	7	13	0.85	0.54
ผู้จัดการมรดก (administrator)	13	8	21	0.62	0.38
ทายาท (heir)	17	10	20	0.85	0.50
ศาล (court)	2	2	3	0.67	0.67
ผู้ตาย (deceased)	3	1	4	0.75	0.25
พยาน (witness)	3	2	4	0.75	0.50
พินัยกรรม (wills)	10	5	13	0.77	0.38
บุตรนอกกฎหมาย (illegitimate child)	2	1	2	1.00	0.50
ถอน (revoke)	2	1	3	0.67	0.33
เจ้าหนี้ (creditor)	3	1	4	0.75	0.25
Average				0.77	0.43

From the experimental results, we found that SCRO algorithm can retrieve a set of diversified documents for user. The average degree of diversity of SCRO is 77% where as the baseline algorithm gets only 43% of diversity level. Moreover, we found that SCRO algorithm has an ability to retrieve more documents for user.

2 The second experiment

We collect 793 Supreme Court sentences related to bill of exchange law. The ontology used for the query expansion are shown in Fig 6.

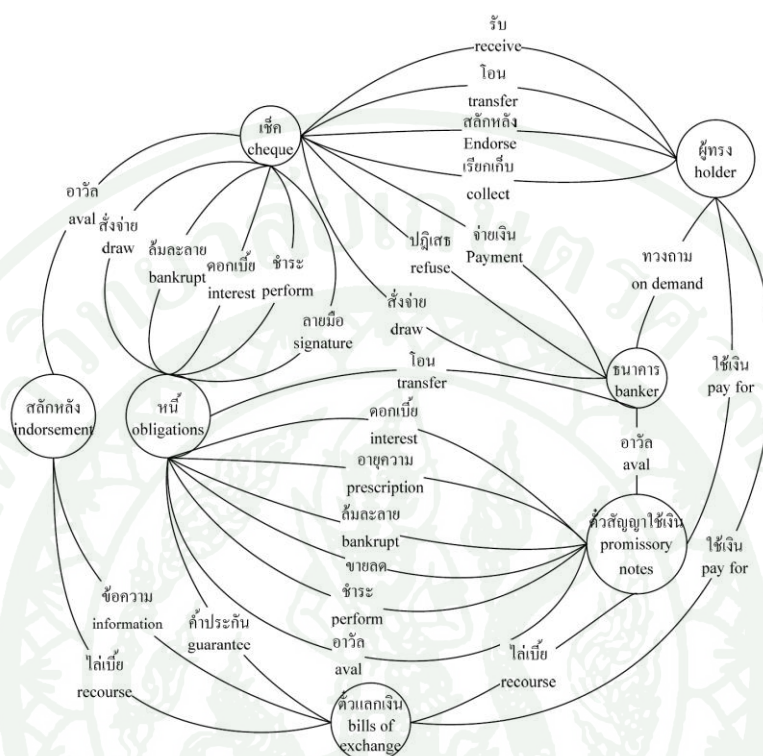


Fig. 6. An ontology for bill of exchange law

We use 13 input keywords to test the retrieval performance which are เช็ค (cheque), ตั๋วแลกเงิน (bills of exchange), ตั๋วสัญญาใช้เงิน (promissory notes), ผู้ทรง (holder), ธนาคาร (banker), หนี้ (obligations), สลักหลัง (endorse), ดอกเบี้ย (interest), โอน (transfer), ไล่เบี้ย (recourse), อาวัล (aval), สั่งจ่าย (draw), ลายมือ (handwriting). The experiments are performed to compare the performance between SCRO algorithm and the baseline algorithm. The experimental results are shown in Table 2.

We found that SCRO algorithm provides more diversity than the baseline algorithm. Considering the average performance, we found that the diversity achieved by SCRO is 71% where as the baseline gets only 58% on diversity level.

Table 2. The performance comparison between SCRO algorithm and the baseline algorithm on bill of exchange data set

keyword	NO. of law issues		Total law issues	Diversity (DC)	
	SCRO	Baseline		SCRO	Baseline
เช็ค (cheque)	10	7	13	0.77	0.54
ตั๋วแลกเงิน (bills of exchange)	3	3	5	0.60	0.60
ตั๋วสัญญาใช้เงิน (promissory notes)	7	7	9	0.78	0.78
ผู้ทรง (holder)	4	4	7	0.57	0.57
ธนาคาร (banker)	5	4	6	0.83	0.67
หนี้ (obligations)	10	5	13	0.77	0.38
สลักหลัง (endorse)	2	2	3	0.67	0.67
ดอกเบี้ย (interest)	2	1	2	1.00	0.50
โอน (transfer)	1	1	2	0.50	0.50
ไล่เบี้ย (recourse)	1	1	2	0.50	0.50
อาวัล (aval)	4	3	4	1.00	0.75
สั่งจ่าย (draw)	1	1	2	0.50	0.50
Average				0.71	0.58

Conclusion

This study proposes an ontology approach for Supreme Court Sentence Retrieval. The objective of using ontology is to provide a set of diversify documents for user. The experiments demonstrate that the proposed ontological approach (SCRO algorithm) outperforms the traditional keyword-based retrieval system.

Reference

- [1] Drid, O. and Ahmed Ben, M. 2008. Building an Ontology-Based Framework For Semantic Information Retrieval: Application To Breast Cancer. In 3rd International Conference.
- [2] Wu, F., Guoshi, W., and Xiangling, Fu. 2007. Design and Implementation of Ontology-Based Query Expansion for Information Retrieval. In Springer Boston. pp.293-298.
- [3] Benjamins, V.R., Contreras, J., Casanovas, P., Ayuso, M., Becue, M., Lemus, L. and Urios, C. 2004. Ontologies of Professional Legal Knowledge as the Basis for Intelligent IT Support for Judges. Artificial Intelligence and Law. Val. 12, pp.359-378.
- [4]. Lee, M., Kun Hua, T., Tzone I, W. 2008. A practical ontology query expansion algorithm for semantic-aware learning objects retrieval. Computers & Education 50. pp.1240-1257.
- [5] Saravanan, M., Ravindran, B. and Raman, S. 2009. Improving legal information retrieval using an ontological framework. In Springer Science Business Media B.V.
- [6] Guiraud, L. and Sylvie, D. 2005. Updating Ontology in the Legal Domain. Proceedings of the 10th international conference on Artificial intelligence and law.

A Study of Thai Succession Law Ontology on Supreme Court Sentences Retrieval

Tanapon Tantisripreecha and Nuanwan Soonthornphisaj

Abstract—This paper presents an improvement of our approach called SCRO_II algorithm. SCRO_I algorithm was initially developed in order to retrieve a set of Supreme Court sentences. The goal of SCRO_I is to provide different law issues among those retrieved documents. We create a new ontology using different semantics to study their performances based on diversity measurement. The contribution of this new ontology is compared to the traditional one. A new procedure is embedded in SCRO_II algorithm to identify a set of synonyms and relations. The experimental results show that using the new ontology as a knowledge representation can enhance the retrieval performance of the algorithm. We study the retrieval performance based on diversity value and found that SCRO_II outperforms SCRO_I algorithm.

Index Terms—Ontology, retrieval, Supreme Court sentences, Thai Succession Law.

I. INTRODUCTION

Decisions made by judges are very important for users in legal fields because these legal documents provide some legal doctrines that can not be found in the code laws. Studying Thai law using only code laws is not possible to fulfill the knowledge of those users. The Supreme Court of Thailand had provided a search facility through the web site (Deka System) (<http://deka2007.supremecourt.or.th/deka/web/search.jsp>). The corpus of the current system consists of 100,010 legal documents. These documents are not predefined into categories of law domain, therefore the retrieval performance is not satisfied in terms of semantic diversity of law content presented in the documents. The reason that legal users need to obtain the variety of law issues from the retrieved document is that they need the complete understanding of each law area in all aspects. To tackle this problem, we propose to use an ontology as a knowledge representation and develop an algorithm called SCRO_I to generate a set of keywords that guarantee the document diversity [1]. However, the performance measured in term of document diversity obtained from SCRO_I algorithm is only 54.6%. The objective of this paper is to improve the

performance of SCRO_I algorithm. We propose a new ontology model that has an ability to enhance the retrieval performance. Another benefit of our proposed algorithm is that users don't need to input several keywords for searching since the algorithm will automatically create a set of keywords for the retrieval process.

This paper is organized as follows. Section 2 describes an overview of ontology and our proposed ontology. Section 3 addresses the SCRO_I and SCRO_II algorithm. Section 4 describes the evaluation details of ontology-based retrieval system. Finally we conclude our work with some discussion and directions of future work in Section 5.

II. ONTOLOGY

A. Related Works on Ontology

Many domain specific ontologies were constructed in various areas such as medical domain [9], e-learning domain [10], legal domain [1]. Ontologies were applied to enhance existing applications such as knowledge-based system [2], computational linguistics [3], etc. Benjamin, *et al.* [6] applied ontology to improve IT support for judges in Spain. They used questionnaires as a tool to collect requirement from users in order to implement the ontology. E-commerce law ontology was constructed using text mining technology by Kayed [7]. A set of conceptual nodes were extracted from e-commerce law cases. They proposed a new algorithm to reduce a number of links connected between nodes in order to decrease the ontology complexity.

Legal Information retrieval research was done for French law [4]. Saravanan, *et al.* did research about query expansion for Indian law in three area which are rent control, income tax and sale tax. They compared their proposed algorithm to the keyword search and found that their algorithm outperformed the traditional retrieval algorithm [5]. The first Korean law ontology was generated using natural language processing technique cooperated with Korean WordNet. The dataset was collected from law text books and research papers about law. After that, their algorithm found a set of word co-occurrences and hand on ability to construct a graph that connects these words together [8].

Manuscript received December 30, 2009.

Nuanwan Soonthornphisaj is with the Department of Computer Science, Faculty of Science, Kasetsart University, Bangkok, Thailand. E-mail: fscinws@ku.ac.th, Phone no. +6689-2007553. (Corresponding Author)

Tanapon Tantisripreecha is with the Department of Computer Science, Faculty of Science, Kasetsart University, Bangkok, Thailand. (e-mail: g5164138@ku.ac.th).

B. Thai Succession Law Ontology

We set up a structure of the Thai succession law ontology using a set of relations and a set of synonyms. Our ontology has two levels, a core level and an extended level ontology.

The core level ontology represents the knowledge of Thai succession law. Fig. 1 shows an example of ontology. In case that user enters keyword, 'r₁', SCRO_II algorithm traverses through the core level ontology and found the relation 'r₁' exists. Next, the algorithm explores all links attached to node between 'r₁', (r₁*, r₂, r₃, r₄). The star sign (*) found in r₁ means that there is an extended knowledge related to r₁. Therefore the algorithm explores the extended level of r₁ and found word1, word2 and word3. Finally a set of queries (Q) are obtained as {A, r₁, B}, {A, word1, B}, {A, word2, B}, {A, word3, B}, {A, r₂, B}, {A, r₃, B} and {A, r₄, C}, respectively.

In case that the keyword obtained from user is found in the node (eg. 'A') of ontology, SCRO_II traverses through the ontology and generates the set of queries as follows: {A, r₁, B}, {A, r₂, B}, {A, r₃, B}, {A, r₄, C}, {A, r₁, S₁}, {A, r₂, S₁}, {A, r₃, S₁}, {A, r₁, S₂}, {A, r₂, S₂}, {A, r₃, S₂}, {A, r₁, S₃}, {A, r₂, S₃}, {A, r₃, S₃}, respectively.

We develop an application for a law expert to create Thai succession law ontology. First, the conceptual nodes were created when user enter keywords into the system. Then the links were connected when user enter words that represent the relationship between nodes. The synonym and reason relation are automatically created using ThaiLegalWordnet.

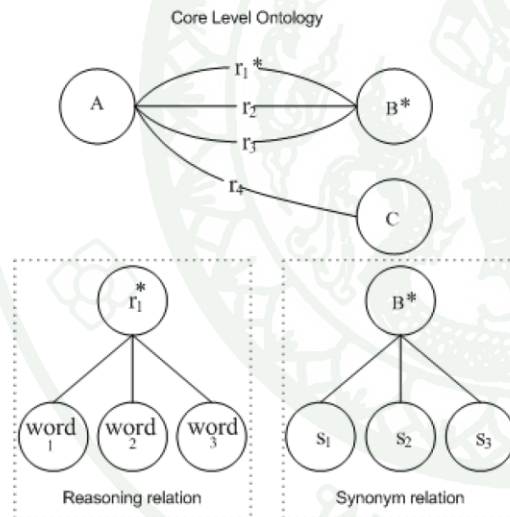


Fig 1. The proposed framework of ontology is divided into core level and extended level.

III. ALGORITHMS

A. SCRO_II algorithm

The framework of SCRO_II algorithm is shown in Fig.2. Given a keyword obtained from user, SCRO_II searches through the core ontology for the keyword. If the word is found in the node of the ontology, the algorithm will traverse through the link to the next node. The words occurred in the nodes and relations are generated to form a set of query.

In case that user enters keyword, 'A', SCRO_I algorithm traverses through the ontology and found the node 'A'. Next, the algorithm explores all links attached with node 'A', (r₁, r₂, r₃, r₄, word1, word2, word3). Then a set of queries (Q) are obtained as {A, r₁, B}, {A, word1, B}, {A, word2, B}, {A, word3, B}, {A, r₂, B}, {A, r₃, B} and {A, r₄, C}, respectively.

SCRO_II also explores the extended level of ontology under two different conditions. In case that the keyword is found in the node, the extended level ontology will be explored only when the star sign is found in the node.

Table 1. SCRO_II algorithm

Algorithm : SCRO_II

Input: Ontology, keyword

Begin

Word = {}

Query = {}

ConceptType = {NODE, RELATION}

SearchConcept (Ontology, keyword)

Word = OntologyTraversal (keyword)

For Each w_i in Word **Do**

If ConceptType(keyword) = NODE and ConceptType(w_i) = NODE

If FoundExtendedOntology(w_i)

ExtendedWord = ExtendedOntologyTraversal()

If ConceptType(w_i) = RELATION and ConceptType(w_i) = RELATION

If FoundExtendedOntology(w_i)

ExtendedWord = ExtendedOntologyTraversal()

End For

Query = QueryGenerate(Word, ExtendedWord)

For each q_i in Query

D_i = DocumentRetrieval (q_i)

Ranking(D_i)

End For

CombineResults = checkDuplicate(D)

End

Output: CombineResults

On the other hand, if the keyword is the relation found in the ontology, the extended ontology will be explored when the star sign is found on that relation. Then the set of queries are created and used in the retrieval process. Considering a set of retrieved documents, the SCRO_II algorithm ranks these documents using TFIDF formula (see equation 1)

$$w_{in} = tf_{in} \times idf_i \quad (1)$$

where w_{in} = weight value of word_i in the retrieved document_n
 tf_{in} = the number of occurrences for word_i in document_n

$$idf = \ln \left(\frac{N}{n_i} \right) \quad (2)$$

where N = the number of retrieved documents
 n_i = the number of retrieved documents contain word_i

SCRO_II algorithm selects the first ranked document as an output for each query and checks for duplicated documents. Since, each query (q_i) is processed independently, therefore the duplicate checking process is done to finalize the output. Because the objective of SCRO_II algorithm is to ensure the diversity of law issues from the final set of retrieved documents.

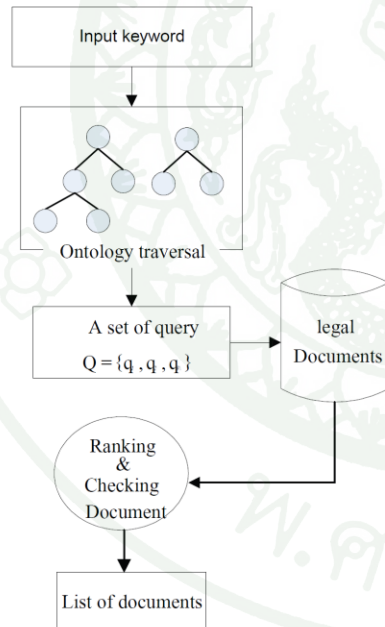


Fig 2. The framework of SCRO_II algorithm.

B. SCRO_I algorithm

The ontology used in SCRO_I algorithm has only one level. Given a keyword, the algorithm explores through the ontology and generates a set of query. The set of query obtained from SCRO_I is different from SCRO_II that yields different performance measure in term of diversity.

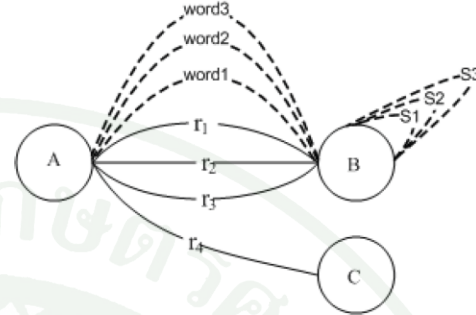


Fig 3. The structure of ontology in SCRO_I

In case that the keyword obtained from user is found in the node 'A', SCRO_II traverses through the ontology and generates the set of queries as follows: $\{A, r_1, B\}$, $\{A, r_2, B\}$, $\{A, r_3, B\}$, $\{A, r_4, C\}$, $\{A, \text{word}_1, B\}$, $\{A, \text{word}_2, B\}$, $\{A, \text{word}_3, B\}$, respectively.

If user searches input r_1 , the set of queries obtained from SCRO_I is $\{A, r_1, B\}$, $\{A, r_2, B\}$, $\{A, r_3, B\}$, $\{A, \text{word}_1, B\}$, $\{A, \text{word}_2, B\}$, $\{A, \text{word}_3, B\}$, respectively.

Table 2. SCRO_I algorithm

Algorithm : SCRO_I
Input: Ontology, keyword
Begin
Word = {}
Query = {}
Word = OntologyTraversal (keyword)
Query = QueryGenerate (Word, Ø)
For each q_i in Query
D_i = DocumentRetrieval (q_i)
Ranking(D_i)
End For
CombineResults = checkDuplicate(D)
End
Output: CombineResults

C. Baseline Algorithm and Deka system

We compare the performance of SCRO_II algorithm with the baseline algorithm and Deka system. The baseline system uses term frequency and inverse document frequency to determine the retrieved document. Whereas the Deka system (Fig. 4) simply uses SQL statements to retrieve legal documents containing a keyword. The current system

implemented by the Thai Supreme Court can be accessed via the website www.deka2007.supremecourt.or.th.

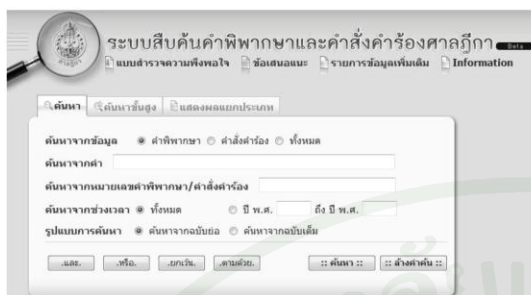


Fig. 4. The Supreme Court Sentences Retrieval System (Deka)

IV. EXPERIMENTS

A. Data set

We collect all judge sentences related to Thai succession law. There are 248 documents in the corpus. The word segmentation is performed since Thai language has no explicit word boundary.

B. Performance measurement

Given a keyword, a set of retrieved documents are measured based on the degree of diversity coverage (DC). The maximum degree of diversity is 1. It means that the algorithm can retrieve all different law issues related to the keyword.

$$DC = \frac{\text{No. of law issues from the set of retrieved documents}}{\text{Total no. of law issues found in ontology}} \quad (3)$$

To test the performance, we compare the retrieval performance of SCRO_II to SCRO_I, baseline algorithm and Deka system, respectively.

In order to evaluate the retrieval performance on Thai succession law, we set up the experiment using 12 keywords which are มรดก (estate), ผู้จัดการมรดก (administrator of an estate), ทายาท (heir), ผู้ตาย (deceased), พยาน (witness), พันຍกรรม (wills), เจ้าหนี้ (creditor), กำจัดมิให้รับมรดก (exclusion for the succession), สละมรดก (renunciation of an estate), พันຍกรรมไม่สมบูรณ์ (incomplete will) ขาระหนี้ (performing loan) ตั้งผู้จัดการมรดก (appointment of the administrator of an estate), respectively.

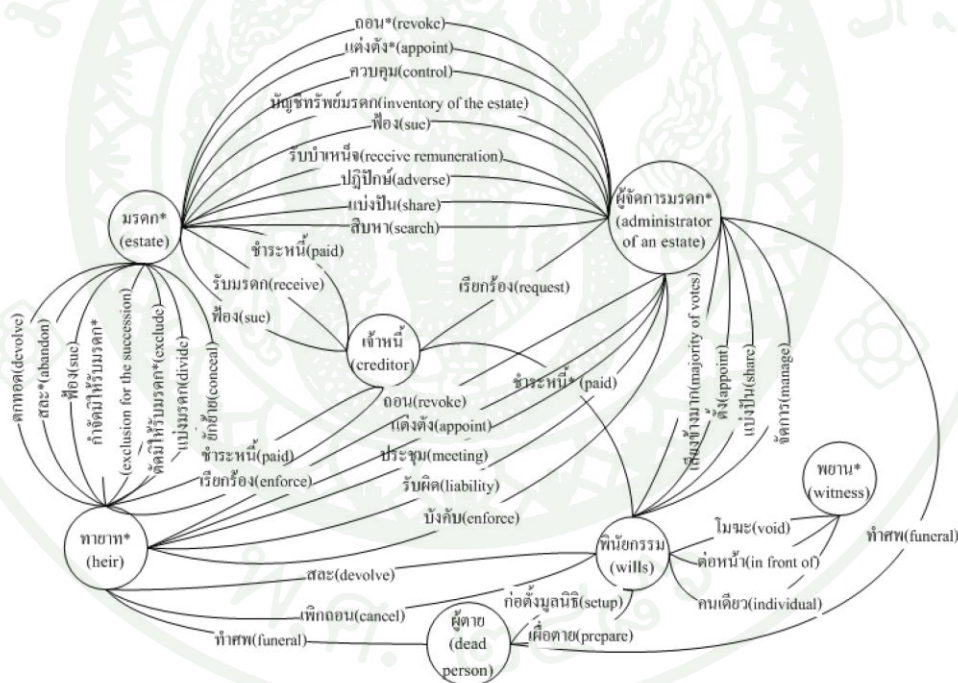


Fig. 3 Thai Succession Law Ontology

Table 3. The comparison of performance in term of the number of law issues found by algorithms

keyword	The number of	The Number of law issues found by algorithms			
	predefined Issues	SCRO II	SCRO I	Baseline	Deka
มรดก (estate)	19	14	8	8	10
ผู้จัดการมรดก (administrator of an estate)	19	10	10	6	6
ทายาท (heir)	17	14	9	9	8
ผู้ตาย (dead person)	4	4	3	1	2
พยาน (witness)	3	3	3	2	1
พินัยกรรม (wills)	9	6	9	2	2
เจ้าหนี้ (creditor)	7	4	3	1	2
ก่าจัดมิให้รับมรดก (exclusion for the succession)	6	4	2	2	6
สละมรดก (renunciation of an estate)	3	2	2	1	3
พินัยกรรมไม่สมบูรณ์ (incomplete will)	4	3	2	2	1
ชำระหนี้ (performing loan)	3	2	1	1	1
ตั้งผู้จัดการมรดก (appointment of the administrator of an estate)	5	3	3	2	1

Table 4. The comparison of performance in term of Diversity measurement found by algorithms

keyword	Diversity Measurements			
	SCRO II	SCRO I	Baseline	Deka
มรดก (estate)	0.74	0.42	0.42	0.53
ผู้จัดการมรดก (administrator of an estate)	0.53	0.53	0.32	0.32
ทายาท (heir)	0.82	0.53	0.53	0.47
ผู้ตาย (dead person)	1.00	0.75	0.25	0.50
พยาน (witness)	1.00	1.00	0.67	0.33
พินัยกรรม (wills)	0.67	1.00	0.22	0.22
เจ้าหนี้ (creditor)	0.57	0.43	0.14	0.29
ก่าจัดมิให้รับมรดก (exclusion for the succession)	0.67	0.33	0.33	1.00
สละมรดก (renunciation of an estate)	0.67	0.67	0.33	1.00
พินัยกรรมไม่สมบูรณ์ (incomplete will)	0.75	0.50	0.50	0.25
ชำระหนี้ (performing loan)	0.67	0.33	0.33	0.33
ตั้งผู้จัดการมรดก (appointment of the administrator of an estate)	0.60	0.60	0.40	0.20
Average	0.72	0.59	0.37	0.45

C. Experimental Results

Considering table 3, we found that SCRO_II can retrieve a set of legal documents with more diversity than those of SCRO_I. The most difficult keyword for algorithms are “estate” and “heir” because they serve as a general term of the domain, so retrieving a set of diversified document for such keyword is considered to be hard. For the keyword, “estate”, SCRO_II is able to retrieve a set of documents that have 14 issues whereas SCRO_I gets only 8 issues. For the keyword, “heir”, SCRO_II gets 14 different issues from a set of retrieved documents whereas SCRO_I gets 9 issues.

Considering Table 4, the performance measured on diversity shows that the average performance of SCRO_II is higher than that of SCRO_I. SCRO_II gets 0.72, whereas SCRO_I gets 0.59.

SCRO_II and SCRO_I can completely retrieve all issues related to the keyword, ผู้ตาย (dead person) and พยาน (witness). Moreover, SCRO_II and SCRO_I have the same performance on keyword, ผู้จัดการมรดก (administrator of an estate) and สละมรดก (renunciation of an estate).

Considering the baseline algorithms, we surprising found that the Deka system using SQL statement outperforms the traditional information retrieval system using tfidf formula. The reason is that Deka system simply retrieves judges sentences having the keyword and sorts by the document ID no. It is assumed that the succession law cases submitted to the Supreme Court are likely to have different issues. However, the Deka system has less diversity value than SCRO_II because the criteria to determine the performance of any information retrieval system concerns about the ranking process. We want the system to deliver the suitable number of choices (documents) to users so that they can effectively view their search results. Therefore, the objective of our research is to develop an algorithm to effectively retrieve a suitable number of judge sentences that guarantee the variety of issues.

V. CONCLUSION

This study proposes an ontological approach for Supreme Court Sentence Retrieval. The objective of using ontology is to provide a set of diversify documents for user. We found that using hierarchical style of ontology can provide more variety of law issues in a set of retrieval results. The experiments demonstrate that SCRO_II outperforms SCRO_I algorithm in term of diversity value. Users are satisfied with the judge documents retrieved by SCRO_II.

ACKNOWLEDGEMENT

This research was partly supported by Faculty of Science and the graduate school of Kasetsart University.

REFERENCES

- [1] T. Tantisriprecha and N. Soonthomphisaj, “Query Expansion Algorithm for Supreme Court Sentences Retrieval using Ontology,” in Proc. of the 48th kasetsart annual conference, 2009.
- [2] A. Valente and J. Breuker, “Making ends meet: conceptual models and ontologies in legal problem solving,” In Proc. Of the XI Brazilian AI Symposium, 1994, pp. 1-15.
- [3] W. Hohfeld, *Fundamental legal conceptions as applied in legal reasoning*. Yale University Press, London.
- [4] L. Guiraud and Sylvie, D. “Updating ontology in the legal domain”, in Proc of the 10th international conference on Artificial intelligence and law, 2005.
- [5] M. Saravanan, B. Ravindran, and S. Raman, “Improving legal information retrieval using an ontological framework” In Springer Science Business Media B.V, 2009.
- [6] V. R. Benjamins, J. Contreras, P. Casanovas, M. Ayuso, Bécue, M., L. Lemus, and C. Urios, (2006) Ontologies of professional legal knowledge as the basic for intelligent IT support for judges. Artificial Intelligence and Law 12: Page(s) 359-378.
- [7] Ahmad Kayed, (2005) Building e-Laws Ontology: New Approach. OTM Workshops, 2005, pp. 826–835.
- [8] H. Soonhee, J. Youngim, Y. Aesun, and H. C.Kwon, “Building Korean Classifier Ontology Based on Korean WordNet”, 2006, pp. 261-268.
- [9] A. Jovic, M. Prcela and D. Gamberger, “Ontologies in Medical Knowledge Representation” in Proc. of the 29th Int. Conf. on Information Technology Interfaces, 2007, pp. 535-540.
- [10] N. Henze, P. Dolog, and W. Nejdl, (2004) “Reasoning and Ontologies for Personalized E-Learning in the Semantic Web”, Educational Technology & Society, 7 (4), pp.82-97.

ประวัติการศึกษา และการทำงาน

ชื่อ –นามสกุล	นายธนพล ตันติศรีปรีชา
วัน เดือน ปี ที่เกิด	21 เมษายน 2529
สถานที่เกิด	สกลนคร
ประวัติการศึกษา	วท.บ. (วิทยาการคอมพิวเตอร์) (เกียรตินิยมอันดับหนึ่ง) มหาวิทยาลัยขอนแก่น
ทุนการศึกษาที่ได้รับ	ทุนผู้ช่วยสอนจากบัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์ (พ.ศ. 2552-2553)