

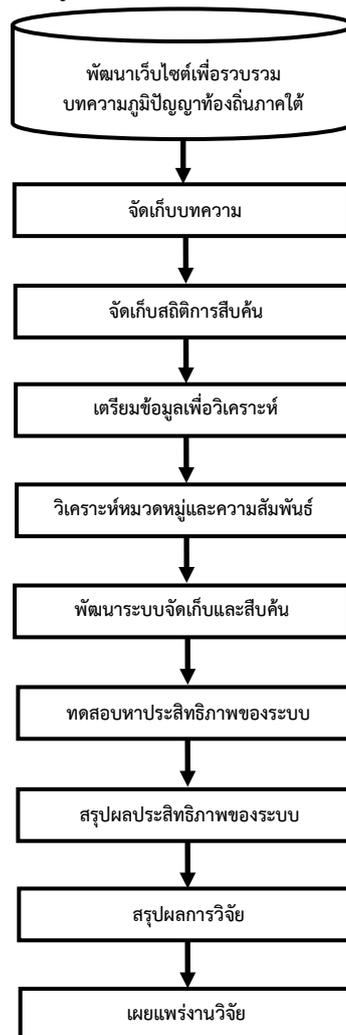
บทที่ 3

วิธีดำเนินการวิจัย

ในการพัฒนาระบบจัดเก็บและสืบค้นภูมิปัญญาท้องถิ่น เพื่อจัดเก็บภูมิปัญญาท้องถิ่น ประกอบด้วยวิธีดำเนินการวิจัย ได้แก่ พัฒนาเว็บไซต์เพื่อรวบรวมบทความภูมิปัญญาท้องถิ่นภาคใต้ จัดเก็บบทความ จัดเก็บสถิติการสืบค้น เตรียมข้อมูลเพื่อวิเคราะห์ วิเคราะห์หมวดหมู่และความสัมพันธ์ พัฒนาระบบจัดเก็บและสืบค้น ทดสอบหาประสิทธิภาพของระบบ สรุปผล ประสิทธิภาพของระบบ สรุปผลการวิจัย และเผยแพร่งานวิจัย

3.1 ขั้นตอนการดำเนินการวิจัย

งานวิจัยนี้ได้นำเทคนิคการทำเหมืองข้อมูลการหาความสัมพันธ์ของบทความและการจัดหมวดหมู่ ในเว็บบล็อกโดยใช้เทคนิคการทำเหมืองข้อมูล โดยได้กำหนดขั้นตอนการดำเนินการวิจัยดังนี้



ภาพที่ 3 - 1 แผนผังขั้นตอนการดำเนินการวิจัย

จากภาพที่ 3-1 สามารถอธิบายแผนผังขั้นตอนการดำเนินการวิจัยได้ดังนี้

3.1.1 สํารวจสืบค้นปราชญ์ท้องถิ่น และลงพื้นที่เก็บข้อมูลองค์ความรู้ เพื่อบรรวมข้อมูล ภูมิปัญญาท้องถิ่นภาคใต้ในพื้นที่ จังหวัดสุราษฎร์ธานี

3.1.2 พัฒนาเว็บไซต์ให้สามารถจัดเก็บบทความ จัดเก็บสถิติการสืบค้น การระบุคำค้นของผู้ใช้ และสถิติการเข้าถึงบทความ จากการค้นโดยตรงในเว็บไซต์ และการสืบค้นผ่านกูเกิ้ล (Google)

3.1.3 นำบทความภูมิปัญญาท้องถิ่นภาคใต้บรรจุลงเว็บไซต์เพื่อเผยแพร่ให้กับเครือข่ายห้องสมุด มหาวิทยาลัยราชภัฏทั่วประเทศให้ เพื่อให้เป็นที่ยุ้จักและได้เข้าสืบค้นข้อมูล ในงานประชุม Workshop on Uninet Network and Computer Application: Wunca ครั้งที่ 24 โครงการ อบรม เมทาเดตา (Metadata) สำหรับการจัดการองค์ความรู้ วันที่ 28-29 กรกฎาคม 2554 ใช้เวลา จัดเก็บข้อมูล 1 ปี 7 เดือน

3.1.4 เตรียมข้อมูลเพื่อวิเคราะห์ ส่งออกข้อมูลสถิติการค้นหาข้อมูล และการเข้าใช้งาน จัดรูปแบบ สถิติการค้นหาข้อมูลของผู้ใช้ เพื่อให้อยู่ในรูปแบบของตัวเลข และตรรกศาสตร์เพื่อ วิเคราะห์ทางสถิติด้วยโปรแกรม Weka

3.1.5 กระบวนการวิเคราะห์ฟอแมลคอนเท็กซ์ (Formal Context) ด้วยโปรแกรม Connexp

3.1.6 สรุปผลหมวดหมู่ และนำหมวดหมู่ที่ได้จากการวิเคราะห์ ไปตรวจสอบโดยผู้ทรงคุณวุฒิ ด้านภูมิปัญญาท้องถิ่นภาคใต้ และสรุปความสัมพันธ์จากคอนเซพแลททิส (Concept Lattice) สำหรับพัฒนาระบบจัดเก็บและสืบค้นภูมิปัญญาท้องถิ่นภาคใต้

3.1.7 จากคอนเซพแลททิส (Concept Lattice) ออกแบบระบบฐานข้อมูลของระบบ ระบบผัง การไหล (DFD) ระบบการจัดเก็บและสืบค้น ประกอบด้วย ระบบการจัดหมวดหมู่บทความ (Cluster Analysis) ระบบจัดทำบทความ ระบบพิจารณาความถี่ของคำ (Frequency Transaction) ระบบ การสร้างความสัมพันธ์ (Association Rule) ออกแบบเชิงแนวคิด (Conceptual Design) เชิง ตรรกะ (Logical Design) ออกแบบเชิงกายภาพ (Physical Design)

3.1.8 พัฒนาระบบจัดเก็บและสืบค้น ระบบการจัดหมวดหมู่บทความ โดยพิจารณาตามความถี่ (Frequency Transaction) ของคำที่ปรากฏเพื่อสร้าง หมวดหมู่เพื่อทำการแบ่งกลุ่มบทความ (Cluster Analysis) พัฒนาระบบความสัมพันธ์ (Association Rule) ของบทความ ด้วยกำหนด เมทาเดตา (Metadata) ให้กับบทความที่มีความเกี่ยวข้องกันของเนื้อหาและคำค้นบางส่วน พัฒนา ระบบการสืบค้นเพื่อค้นหาบทความ ที่มีความเกี่ยวข้องกันจากคำค้นเดียวกัน พัฒนาระบบจัดเก็บและ สืบค้น

3.1.9 ทดสอบหาประสิทธิภาพของระบบ เป็นการประเมินค่า Recall และ Precision เป็นมาตร วัดการเรียกคืนและความถูกต้องแม่นยำวิธีการประเมินประสิทธิภาพระบบจัดเก็บและสืบค้น ซึ่งเป็น วิธีการที่ใช้ในงานวิจัยของ จากงานวิจัยของ สถาพร สังข์ศิริ วาทีณี ดวงอ่อนนาม และธนากร สุวรรณวัฒน์

3.1.10 สรุปกระบวนการวิจัยและผลการวิจัย เขียนบทความตีพิมพ์เผยแพร่ในวารสารที่เป็นที่ ยอมรับ

3.2 พัฒนาเว็บไซต์เพื่อรวบรวมบทความภูมิปัญญาท้องถิ่น

3.2.1 เว็บไซต์ที่พัฒนาขึ้นเป็นโปรแกรมซึ่งเขียนด้วยภาษาPHP และใช้ฐานข้อมูล MySQL มีกระบวนการทำงานและการจัดการข้อมูลในรูปแบบเว็บบล็อก

3.2.2 สามารถเรียงลำดับบทความตามเวลาในการเขียน

3.2.3 สามารถแสดงข้อมูลที่เขียนล่าสุดไว้แรกสุด

3.2.4 สามารถเขียนข้อความ อัปโหลดภาพ ลิงก์ รวมทั้งสื่อ เพลง หรือวิดีโอในหลายรูปแบบได้

3.2.5 ผู้ใช้สามารถอ่านข้อมูล สามารถแสดงความคิดเห็นต่อท้ายข้อความที่เจ้าของบล็อกเป็นผู้เขียน และผู้เขียนสามารถได้ผลตอบกลับโดยทันที

3.2.6 กำหนดช่องทางในการสืบค้นข้อมูลผ่านทางเว็บบล็อก

3.2.7 สามารถจัดเก็บสถิติการเข้าใช้ การระบุค่าค้นของผู้ใช้ และการนำเสนอบทความที่เกี่ยวข้องกับคำค้นที่ผู้ใช้สืบค้น

3.2.8 ประชาสัมพันธ์ และเผยแพร่ต้นแบบเว็บบล็อกธนาคารสมองภูมิปัญญาท้องถิ่น ให้กับเครือข่ายของสมาคมมหาวิทยาลัยราชภัฏทั่วประเทศให้ เพื่อให้เป็นที่รู้จักและได้เข้าสืบค้นข้อมูล ในงานประชุม Workshop on Uninet Network and Computer Application: Wunca ครั้งที่ 24 โครงการอบรม เมทาเดตา (Meta Data) สำหรับ การจัดการองค์ความรู้ ดังแสดงไว้ในภาพที่ 3-2



ภาพที่ 3 - 2 ระบบจัดเก็บและสืบค้น

จากภาพที่ 3-2 สามารถอธิบายได้ว่า เว็บบล็อกใช้ในการรวบรวมบทความองค์ความรู้ภูมิปัญญาท้องถิ่น ประชาสัมพันธ์ในพื้นที่ จ.สุราษฎร์ธานี และพื้นที่ใกล้เคียงรอยต่อจังหวัดสุราษฎร์ธานี รวบรวมบทความเกี่ยวกับภูมิปัญญาท้องถิ่นภาคใต้ เพื่อให้ผู้ใช้งานได้ทดลองใช้งานสืบค้นเกี่ยวกับภูมิปัญญาท้องถิ่นภาคใต้ และเก็บค่าสถิติการเข้าใช้งานจากผู้ใช้ เผยแพร่ต้นแบบเว็บบล็อกธนาคารสมองภูมิปัญญาท้องถิ่น ให้กับเครือข่ายของสมาคมมหาวิทยาลัยราชภัฏทั่วประเทศให้ เพื่อให้เป็นที่รู้จัก

และได้เข้าสืบค้นข้อมูล ในงานประชุม Workshop on Uninet Network and Computer Application: Wunca ครั้งที่ 24 ในโครงการอบรม เมทาเดตาสำหรับ การจัดการองค์ความรู้ ระหว่างวันที่ 28-29 กรกฎาคม 2555 สามารถทดลองเข้าใช้ได้ที่ <http://r-linet.sru.ac.th/KM/>

3.2.9 ออกแบบและพัฒนาระบบจัดเก็บและสืบค้น จากผลการวิเคราะห์ที่ได้วิเคราะห์ ได้แก่ การแบ่งกลุ่มข้อมูล (Cluster Analysis) ของบทความ ความสัมพันธ์ (Association Rule) ของการค้น ผลการวิเคราะห์ฟอแมลคอนเท็กซ์ (Formal Context) ซึ่งแสดงไว้ในบทที่ 4

3.3 จัดเก็บบทความ

3.3.1 จัดทำโครงการเพื่อนำองค์ความรู้เข้าสู่ระบบในการจัดเก็บองค์ความรู้

3.3.2 สืบค้นข้อมูลจากฐานข้อมูลท้องถิ่น ในพื้นที่ จ.สุราษฎร์ธานี และพื้นที่ใกล้เคียงรอยต่อจังหวัดสุราษฎร์ธานี

3.3.3 ติดต่อประสานงานหน่วยงานปกครองส่วนท้องถิ่น เช่น องค์การบริหารส่วนตำบล เทศบาล สำนักงานเขตพื้นที่การศึกษา เพื่อสำรวจรายชื่อท้องถิ่น ในพื้นที่ จ.สุราษฎร์ธานี และพื้นที่ใกล้เคียงรอยต่อจังหวัดสุราษฎร์ธานี

3.3.4 ติดต่อประสานงานจากหน่วยงานท้องถิ่น โดยการทำหนังสือเชิญเพื่อเข้าร่วมโครงการ

3.3.5 เนื่องจากรายชื่อท้องถิ่นท้องถิ่น เป็นผู้สูงอายุ ทางผู้ดูแลจึงไม่สะดวกที่จะให้เดิน และบางท่านก็ได้เสียชีวิตระหว่างโครงการวิจัยยังดำเนินการอยู่ จึงไม่สะดวกที่จะมาถ่ายทอดความรู้ได้ ผู้วิจัยจึงเปลี่ยนแผนงาน เป็นการลงพื้นที่เก็บข้อมูล จากรายชื่อท้องถิ่นท้องถิ่นโดยตรง

3.3.6 ออกหนังสือจากทางคณะวิทยาศาสตร์และเทคโนโลยี เพื่อขอความอนุเคราะห์ให้ถ่ายทอดความรู้เกี่ยวกับภูมิปัญญาท้องถิ่น

3.3.7 ลงพื้นที่เก็บข้อมูลองค์ความรู้ เช่น บทความ สื่อวีดิทัศน์ ภาพถ่าย

3.3.8 บันทึกบทความพร้อมสื่อลงฐานข้อมูล

3.3.9 ออกเกียรติบัตรเชิดภูมิปัญญาท้องถิ่น ให้แก่ ประชาชนท้องถิ่นท้องถิ่น โดยสถาบันวิจัยมหาวิทยาลัยราชภัฏสุราษฎร์ธานี

3.3.10 จัดประชาสัมพันธ์ และเผยแพร่เว็บบล็อกธนาคารสมองภูมิปัญญาท้องถิ่น

3.3.11 ใช้เวลาจัดเก็บข้อมูล 1 ปี 7 เดือน

3.4 เตรียมข้อมูลเพื่อวิเคราะห์

ผู้วิจัยได้จำแนกกรณีการค้นข้อมูล ใช้กับกรณี 1 กลุ่มไอพีแอดเดรส (IP Address) คือ 1 กลุ่มผู้ใช้งาน และกำหนดจำนวนกลุ่ม ตัวแปรที่ใช้จะต้องเป็นตัวแปรเชิงปริมาณ เป็น สเกลอันตรภาค (Interval Scale) สเกลอัตราส่วน (Ratio Scale) ในส่วนของการวิเคราะห์ความสัมพันธ์ (Association Rule) ได้สร้างกฎความสัมพันธ์ระหว่างคำค้นในแต่ละบทความ มีพื้นฐานการศึกษาจากการค้นหาของกรณี 1 กลุ่มไอพีแอดเดรส (IP Address) สำหรับผู้ใช้สนใจค้นหา 245 บทความแบบต่อเนื่องเกิดขึ้นร่วมกันหรือพร้อมกันในฐานข้อมูล เตรียมข้อมูลเพื่อวิเคราะห์ ส่งออกข้อมูลสถิติการค้นหาข้อมูลจาก 7,556 รายการ ทำการซ่อมแซมข้อมูล (Prepare Data) โดยใช้ค่าเฉลี่ยของตัวแทนกลุ่มไปเติมในข้อมูลที่ว่าง และค่าเฉลี่ยของตัวแทนกลุ่มทั้งหมด เพื่อนำข้อมูลการเข้าใช้งาน ซึ่ง

จัดรูปแบบ สถิติการค้นหาข้อมูลของผู้ใช้ เพื่อให้อยู่ในรูปแบบของตัวเลข และตรรกศาสตร์เพื่อวิเคราะห์ทางสถิติด้วยโปรแกรม Weka

3.4.1 การแบ่งกลุ่มข้อมูล (Cluster Analysis) โดยใช้เทคนิค K-means Cluster

3.4.2 ส่งออกข้อมูลสถิติการใช้เว็บบล็อก ในการแสดงการระบุค่าคั่นของผู้ใช้ และการนำเสนอ บทความที่เกี่ยวข้อง

3.4.3 รวบรวมข้อมูลการค้นหาข้อมูลของผู้ใช้จากแฟ้มข้อมูลในคอมพิวเตอร์

3.4.4 จัดรูปแบบข้อมูลการค้นหาข้อมูลของผู้ใช้ให้อยู่ในรูปแบบตัวแปรเชิงตัวเลขเพื่อให้สามารถวิเคราะห์ด้วยโปรแกรม Weka ได้

3.4.5 ปรับข้อมูลให้เป็นรูปแบบของตัวเลข และตรรกศาสตร์เพื่อทำการวิเคราะห์ทางสถิติ เพื่อหาค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน

3.4.6 พิจารณาผลการวิเคราะห์ K-means Cluster และ Hierarchical Cluster เพื่อคัดเลือกผลการวิเคราะห์ที่นำมาใช้ในการวิเคราะห์ การหาค่าเฉลี่ยแบบ K-means Cluster อาจจะได้กลุ่มที่หนาแน่นกว่าการจัดกลุ่มแบบ Hierarchical ข้อมูลการค้นหาข้อมูลของผู้ใช้ จะเกาะกลุ่มกันเป็นกลุ่มมากกว่าระยะห่างจุดในข้อมูลน้อยกว่าจึงน่าเชื่อถือมากกว่า

3.4.7 กำหนดจำนวนกลุ่ม K กลุ่ม และกำหนดจุดศูนย์กลางเริ่มต้นจำนวน K จุด

3.4.8 นำข้อมูลการค้นหาข้อมูลของผู้ใช้ทั้งหมดจัดเข้ากลุ่มที่มีจุดศูนย์กลางที่อยู่ใกล้ศูนย์กลางมากที่สุด โดยคำนวณจากการวัดระยะห่างระหว่างจุดที่น้อยที่สุด

3.4.9 กำหนดหมวดหมู่กลุ่มใหญ่ จากความถี่ที่ปรากฏสูงสุด โดยการเขียนโปรแกรมนับค่าคั่น

3.4.10 จัดรูปข้อมูลหรือตัวแปรเป็นข้อมูลชนิดที่สามารถใช้เทคนิค K-means Cluster

3.4.11 นำเทคนิค K-mean Cluster จัดกลุ่มหมวดหมู่ของบทความจากสถิติการเข้าใช้

3.4.12 สรุปผลการวิเคราะห์ข้อมูล จัดกลุ่มหมวดหมู่ของบทความจากสถิติการเข้าใช้

3.4.13 ประเมินและวัดประสิทธิภาพจากค่า Sum of square error (SSE) นั่นคือ ค่าระยะห่าง (distance) ที่ต่างจากจุดศูนย์กลาง (Center) ของแต่ละคลัสเตอร์

3.4.14 แปรค่าทางสถิติเป็น หมวดหมู่หลัก หมวดหมู่รอง สำหรับพัฒนาระบบจัดเก็บและสืบค้น ภูมิปัญญาท้องถิ่นภาคใต้

3.4.15 นำหมวดหมู่ที่ได้จากการวิเคราะห์ ไปตรวจสอบโดยผู้ทรงคุณวุฒิด้านภูมิปัญญาท้องถิ่น เพื่อประเมินความเป็นไปได้ในการใช้งาน

3.5 วิเคราะห์หมวดหมู่

3.5.1 วิเคราะห์หมวดหมู่ภูมิปัญญาท้องถิ่นภาคใต้จากสถิติการสืบค้น ด้วยอัลกอริทึมเคมีน คลัสเตอร์ริง (K-means Clustering Algorithm) โดยประยุกต์กระบวนการวิเคราะห์มาจากการวิจัยของ ชิตชนก ส่งศิริ ดังนี้

3.5.1.1 Pre-Processing ขั้นตอนการเตรียมข้อมูลให้เหมาะสมและให้อยู่ในรูปแบบที่สามารถนำไปใช้งานได้ โดยกรณี 1 กลุ่ม ไอพีแอดเดรส (IP address) คือ 1 กลุ่มผู้ใช้งานในแต่ละบทความจะมีตัวแทน (Agent) ของบทความ คิดจาก คำคั่นที่มีความถี่สูงเป็นตัวแทนตัวแทน (Agent) ของเนื้อหาบทความการใช้งานของ 1 กลุ่มไอพีแอดเดรส (IP Address) ความถี่ของ 1 กลุ่มไอพี

แอดเดรส (IP Address) แทนการใช้ของ 1 กลุ่มผู้ใช้ ความถี่ของการอ่านบทความ คือ การใช้งานตัวแทน (Agent) ของบทความ

3.5.1.2 ขั้นตอน Data Mining เป็นการพิจารณาระหว่าง K-means Clustering Algorithm กับ Hierarchical Cluster Algorithm เพื่อคัดเลือกอัลกอริทึมที่นำมาใช้ในการวิเคราะห์ โดยเลือกใช้ K-means Clustering Algorithm เพราะรายการข้อมูล (Data Items) ที่ปริมาณมาก มีจำนวน 7,481 รายการ จึงเหมาะสมกว่า Hierarchical Cluster Algorithm ประกอบกับรายการข้อมูล (Data Items) ในแต่ละรายการ ไม่ได้บอกว่ายู่กลุ่มใด เป็นผลจากการจัดกลุ่มของ K-means Clustering Algorithm จากการหาค่าเฉลี่ยแบบ K-means Cluster การกำหนดจำนวนกลุ่ม K กลุ่ม และการกำหนดจุดศูนย์กลางเริ่มต้นจำนวน K จุด

3.5.1.3 Post-processing กำหนดจำนวนกลุ่ม แปลงข้อมูลจากไฟล์ เอสคิวแอล (sql) ให้อยู่ในรูปแบบไฟล์ซีเอสวี (csv) เพื่อวิเคราะห์โดยโปรแกรม Weka โดยทดลองกำหนดกลุ่ม K เพื่อการวิเคราะห์ คือ กรณีการแบ่งกลุ่ม เป็น 10 หมวดหมู่ ค่า Mean Square Error ต่ำสุด เท่ากับ .123 และ สูงสุด เท่ากับ .181 ค่า Sig เท่ากับ .000 สำหรับ กรณีการแบ่งกลุ่ม เป็น 15 หมวดหมู่ ค่า Mean Square Error ต่ำสุด เท่ากับ .033 และ สูงสุด เท่ากับ .143 ค่า Sig เท่ากับ .000 สำหรับ กรณีการแบ่งกลุ่ม เป็น 20 หมวดหมู่ ค่า Mean Square Error ต่ำสุด เท่ากับ .000 และ สูงสุด เท่ากับ .143 ค่า Sig เท่ากับ .020 ในส่วนของประเมินและวัดประสิทธิภาพจำนวนหมวดหมู่หลัก วัดจากค่าผลรวมของกำลังสองของความคลาดเคลื่อน (Sum of square error: SSE) ที่ใกล้เคียงค่า 0 ที่สุด และค่าระยะห่าง (Distance) ที่ต่างจากจุดศูนย์กลาง (Center) ของแต่ละคลัสเตอร์ เพื่อวัดการกระจุกตัวของการแบ่งกลุ่มตลอดจนแปรค่าทางสถิติเป็นจำนวนหมวดหมู่หลัก (ชิตชนก ส่งศิริ, 2544)

3.6 วิเคราะห์ความสัมพันธ์

งานวิจัยนี้เลือกใช้เทคนิคการค้นหากฎความสัมพันธ์ (Association Rule) โดยใช้เทคนิค Apriori ซึ่งศึกษาจากงานวิจัยของ Mohan, E.K. & Han และ Agrawal, Rakesh, & Ramakrishnan Srikant เพื่อใช้ในการคาดการณ์ความสนใจบทความภูมิปัญญาท้องถิ่น โดยพิจารณาจากร้อยละของค่าความเชื่อมั่น และค่าสนับสนุนที่เกิดขึ้น สามารถหาค่าความเชื่อมั่น และค่าสนับสนุนจากสมการดังนี้

$$\text{ค่าความเชื่อมั่น (A,B)} = \frac{\text{จำนวนของ Transaction (A,B)}}{\text{จำนวน Transaction (A)}} \quad (3-1)$$

$$\text{ค่าสนับสนุน (A,B)} = \frac{\text{จำนวนของ Transaction (A,B)}}{\text{จำนวน Transaction ทั้งหมด}} \quad (3-2)$$

โดยที่ A หมายถึง เหตุการณ์ที่ใช้เป็นเงื่อนไขในการคาดการณ์

B หมายถึง เหตุการณ์ที่เป็นผลลัพธ์ที่ได้จากการคาดการณ์

Transaction (A, B) หมายถึง เหตุการณ์ที่ประกอบด้วยเหตุการณ์ A และ B

Transaction (A) หมายถึง เหตุการณ์ที่ประกอบด้วยเหตุการณ์ A อย่างเดียว

การหาความสัมพันธ์ทั้งหมดจะต้องมีค่าสนับสนุนมากกว่าค่าสนับสนุนต่ำสุดที่กำหนดไว้ และมีค่าความเชื่อมั่นมากกว่าค่าความเชื่อมั่นต่ำสุดที่กำหนดไว้ ในส่วนของการวิเคราะห์ความสัมพันธ์ (Association Rule) จากการรวบรวมข้อมูลผลการค้นหาข้อมูลแบบต่อเนื่อง จำนวน 57 รายการที่สัมพันธ์กัน ไม่เพียงพอที่จะวิเคราะห์ความสัมพันธ์ มีปริมาณที่น้อยเกินไป สำหรับการทำให้เหมือนข้อมูล ส่งผลให้ขาดความน่าเชื่อถือของผลลัพธ์ จึงเลือกวิเคราะห์ความสัมพันธ์ (Association Rule) โดยใช้การวิเคราะห์ฟอแมลคอนเท็กซ์ (Formal Context) (Mohan, E.K. & Han, 2011; Agrawal, Rakesh, & Ramakrishnan Srikant, 1994)

3.7 กระบวนการวิเคราะห์ฟอแมลคอนเท็กซ์

จากทฤษฎีการแบ่งกลุ่มข้อมูล (Cluster Analysis) และทฤษฎีการหาความสัมพันธ์ (Association Rule) นำผลข้อมูลที่ได้จากการวิเคราะห์มาเปรียบเทียบกับผลการวิเคราะห์บทความ โดยใช้การวิเคราะห์ฟอแมลคอนเท็กซ์ (Formal Context) ใช้กระบวนการโดยมีขั้นตอนดังนี้

3.7.1 การวิเคราะห์บทความ

นำบทความที่รวบรวมจากการลงพื้นที่ จัดเก็บในรูปแบบแท็กไฟล์ นามสกุล txt เพื่อทำแยกคำ ซึ่งเป็นวิธีการหนึ่งที่ใช้ในงานวิจัยของ อธิวัฒน์ พงษ์ศิริปริดา แต่แตกต่างตรงที่ งานวิจัยชิ้นนี้ เลือกใช้อัลกอริทึม Swath ของ ไพศาล เจริญพรสวัสดิ์ ในการแยกคำต่าง ๆ ในบทความ และบันทึกไฟล์สำหรับการวิเคราะห์ด้วยโปรแกรมนับคำ โดยไฟล์ที่ได้จากการตัดคำด้วยอัลกอริทึม Swath จะแยกคำต่าง ๆ ในบทความมีการตัดคำแยกด้วยตัวคั่น “|” เพื่อให้สามารถนำไฟล์ที่ได้ ไปนับคำด้วยโปรแกรม

3.7.1.1 การตัดคำในบทความ โดยใช้ อัลกอริทึม Swath

ผู้วิจัยนำบทความที่รวบรวมจากการลงพื้นที่ จัดเก็บในรูปแบบแท็กไฟล์ นามสกุล txt เพื่อทำการวิเคราะห์ด้วย อัลกอริทึม Swath ของ ไพศาล เจริญพรสวัสดิ์ เพื่อแยกคำต่าง ๆ ในบทความ และบันทึกไฟล์การวิเคราะห์สำหรับการเขียนโปรแกรมในการนับคำ

```
D:\>cd\swath
D:\swath>swath.exe -h
USAGE: swath [mule|-v] [-b "wordseparator"] [-d wordsegdatadir] [-f html|rtf|latex|lambda|winlatex|maclatex] [-m long|max|bi|bip] [-l] [-help]
D:\swath>swath.exe -b "|" -d data <test01.txt> out.txt
D:\swath>
```

ภาพที่ 3-3 กระบวนการตัดคำในบทความ โดยใช้ อัลกอริทึม Swath

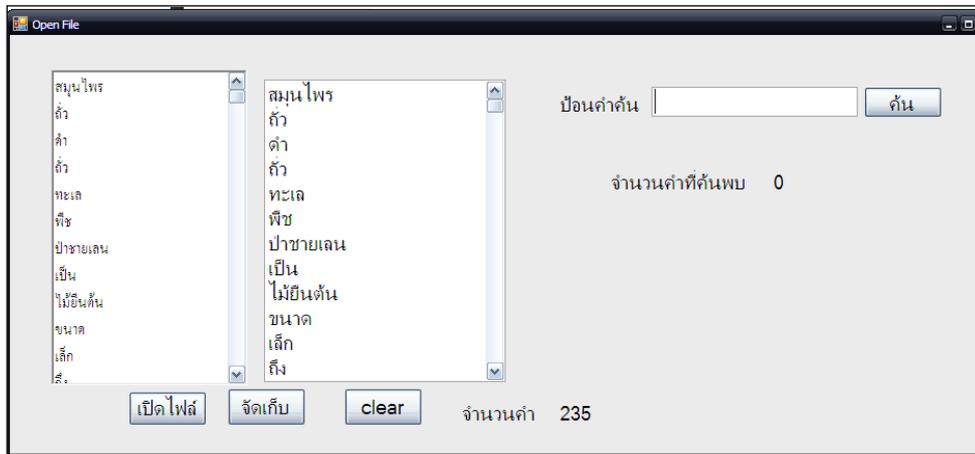
การตัดคำด้วยอัลกอริทึม Swath แยกคำต่าง ๆ ในบทความจะมีการตัดคำแยกด้วยตัวคั่น “|” เพื่อให้สามารถนำไฟล์ที่ได้ ไปเขียนเป็นโปรแกรม เพื่อใช้ในการนับคำ ซึ่งผลลัพธ์การตัดคำที่ได้จะแสดงในภาพที่ 3-4

สมุนไพรร
 ถั่ว|ดำ| |ถั่ว|ทะเล| |พืช|ป่าชายเลน| |เป็น|ไม้ยืนต้น|ขนาด|เล็ก|ถึง|กลาง| |พื|พอ|น|น้อย|แต่|บริเวณ|โคน| |ผิว|เปลือก
 สี|น้ำตาล|เขียว| |เมื่อ|แก่|ออก|ดอก|และ|ผล|เกือบ|ตลอด|ปี| |การ|เจริญ|เติบโต|ขึ้น|ใน|พื้นที่|ด้าน|ใน|ของ|ป่าชายเลน
 ถั่ว|ดำ|ใช้|รักษา|โรค|ทางเดินอาหาร|อีก|เสมอ|ใน|เด็ก|

ภาพที่ 3-4 ผลลัพธ์การตัดคำ โดยใช้ อัลกอริทึม Swath

3.7.1.2 การเขียนโปรแกรมเพื่อนับคำในบทความ

ผู้วิจัยและคณะได้เขียนโปรแกรมช่วยนับคำในบทความ สำหรับการวิเคราะห์คำค้นในบทความ และนับความถี่ของคำค้น เพื่อจัดบทความเข้าสู่หมวดหมู่ที่ได้วิเคราะห์ไว้ ดังแสดงไว้ในภาพที่ 3-5



ภาพที่ 3-5 โปรแกรมเพื่อนับคำในบทความ

3.7.2 พอมอลคอนเท็กซ์ (Formal Context)

ชนวิวัฒน์ ศรีสอ้าน (2551) ได้กำหนดรูปแบบตารางประกอบไปด้วย 3 ส่วน คือ ออบเจกต์ (Object) แอตทริบิว (Attribute) ความสัมพันธ์ (Relation) ทฤษฎี พอมอลคอนเท็กซ์ (Formal Context) ประกอบด้วย 2 ส่วนได้แก่ ขอบเขต (extent) หรือ เซต (Set) ของวัตถุในตารางที่เกิดร่วมกัน คือวัตถุในตาราง และ เจตนา (Intent) คือลักษณะแอตทริบิว (attribute) ร่วมกันของวัตถุ การนิยามตัวแปรต่างๆดังนี้

O = ออบเจกต์ (Object) A = แอตทริบิว (Attribute) R = ความสัมพันธ์ (Relation)
 กำหนดให้ $X \subseteq O$ และ $Y \subseteq A$

สมการ $P(X) = \{a \in A \mid \forall o(X: (o, a)(R))\}$ หมายถึง แอตทริบิวร่วมของ X (3-3)

สมการ $Q(X) = \{a \in A \mid \forall o(Y: (o, a)(R))\}$ หมายถึง แอตทริบิวร่วมของ Y (3-4)

สามารถอธิบายได้ว่า Context K: = (O, A, R) ออบเจกต์ (Object) ประกอบด้วย
 O:= {01 สมุนไพรร, 02 ยารักษาโรค, 03 ผักและอาหารพื้นบ้าน, 04 การเลี้ยงเด็ก, 05 เพลงกล่อมเด็ก, 06 นิทานพื้นบ้าน, 07 สุภาษิต คำพังเพย, 08 ศาสนา, 09 ประเพณีวัฒนธรรม, 10 คาถา, 11 หัตถกรรม, 12 เครื่องมือเครื่องใช้การดำรงชีวิต, 13 ดนตรี, 14 การละเล่น, 15 การแสดง}

ตารางที่ 3-3 วิเคราะห์ ฟอมอลคอนเท็กซ์ (Formal Context) ของบทความในหมวด 03 ผักและอาหารพื้นบ้าน

		C0001	C0002	C0003	C0004	C0005	C0006	C0007	C0008	C0009	C0010	C0011	C0012	C0013	C0014	C0015	C0016	C0017	C0018	C0019	C0020	C0021	C0022	C0023	C0024	C0025	C0026	C0027	C0028	C0029
01	สมุนไพร	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0
02	ยารักษาโรค	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
03	ผักและอาหารพื้นบ้าน	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
04	การเลี้ยงเด็ก	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
05	เพลงกล่อมเด็ก	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
06	นิทานพื้นบ้าน	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
07	สุภาอิต คำพังเพย	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
08	ศาสนา	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
09	ประเพณีวัฒนธรรม	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	คาถา	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	หัตถกรรม	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	การดำรงชีวิต	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	1	0	1	1	0	0	0
13	ดนตรี	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	การละเล่น	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	การแสดง	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ตารางที่ 3-4 วิเคราะห์ ฟอมอลคอนเท็กซ์ (Formal Context) ของบทความในหมวด 04 การเลี้ยงเด็ก

		D0001	D0002	D0003
01	สมุนไพร	1	1	0
02	ยารักษาโรค	1	0	0
03	ผักและอาหารพื้นบ้าน	1	1	0
04	การเลี้ยงเด็ก	1	1	1
05	เพลงกล่อมเด็ก	0	0	0
06	นิทานพื้นบ้าน	0	0	0
07	สุภาอิต คำพังเพย	0	0	0
08	ศาสนา	0	0	0
09	ประเพณีวัฒนธรรม	0	0	0
10	คาถา	0	0	0
11	หัตถกรรม	0	0	0
12	การดำรงชีวิต	0	0	1
13	ดนตรี	0	0	0
14	การละเล่น	0	0	0
15	การแสดง	0	0	0

ตารางที่ 3-7 วิเคราะห์ ฟอมอลคอนเท็กซ์ (Formal Context) ของบทความในหมวด 07 สุภาชิต คำพังเพย

		G0001	G0002	G0003	G0004	G0005	G0006	G0007	G0008	G0009	G0010	G0011	G0012
01	สมุนไพรม	0	0	0	0	0	0	0	0	0	0	0	0
02	ยารักษาโรค	0	0	0	0	0	0	0	0	0	0	0	0
03	ผักและอาหารพื้นบ้าน	0	0	0	0	0	0	0	0	0	0	0	0
04	การเลี้ยงเด็ก	0	0	0	0	0	0	0	0	0	0	0	0
05	เพลงกล่อมเด็ก	0	0	0	0	0	0	0	0	0	0	0	0
06	นิทานพื้นบ้าน	0	0	0	0	0	0	0	0	0	0	0	0
07	สุภาชิต คำพังเพย	1	1	1	1	1	1	1	1	1	1	1	1
08	ศาสนา	0	0	0	0	0	0	0	0	0	0	0	0
09	ประเพณีวัฒนธรรม	0	0	0	1	0	0	0	0	0	0	0	0
10	กาลา	0	0	0	0	0	0	0	0	0	0	0	0
11	หัตถกรรม	0	0	0	0	0	0	0	0	0	0	0	0
12	การดำรงชีวิต	0	0	0	1	0	0	0	0	0	0	0	0
13	ดนตรี	0	0	0	0	0	0	0	0	0	0	0	0
14	การละเล่น	0	0	0	0	0	0	0	0	0	0	0	0
15	การแสดง	0	0	0	0	0	0	0	0	0	0	0	0

ตารางที่ 3-8 วิเคราะห์ ฟอมอลคอนเท็กซ์ (Formal Context) ของบทความในหมวด 08 ศาสนา

		H0001	H0002	H0003	H0004	H0005	H0006
01	สมุนไพรม	0	0	0	0	0	0
02	ยารักษาโรค	0	0	0	0	0	0
03	ผักและอาหารพื้นบ้าน	0	0	0	0	0	0
04	การเลี้ยงเด็ก	0	0	0	0	0	0
05	เพลงกล่อมเด็ก	0	0	0	0	0	0
06	นิทานพื้นบ้าน	0	0	0	0	0	0
07	สุภาชิต คำพังเพย	0	0	0	0	0	0
08	ศาสนา	1	1	1	1	1	1
09	ประเพณีวัฒนธรรม	1	1	1	1	1	1
10	กาลา	0	0	0	0	0	0
11	หัตถกรรม	0	0	0	0	0	0
12	การดำรงชีวิต	0	1	1	0	1	1
13	ดนตรี	0	0	0	0	0	0
14	การละเล่น	0	0	0	0	0	0
15	การแสดง	0	0	0	0	0	0

ตารางที่ 3-13 วิเคราะห์ฟอมอลคอนเท็กซ์ (Formal Context) ของบทความในหมวด 13 ดนตรี

		M0001	M0002	M0003	M0004	M0005	M0006	M0007	M0008
01	สมุนไพรมะ	0	0	0	0	0	0	0	0
02	ยารักษาโรค	0	0	0	0	0	0	0	0
03	ผักและอาหารพื้นบ้าน	0	0	0	0	0	0	0	0
04	การเลี้ยงเด็ก	0	0	0	0	0	0	0	0
05	เพลงกล่อมเด็ก	0	0	0	0	0	0	0	0
06	นิทานพื้นบ้าน	0	0	0	0	0	0	0	0
07	สุภาอิต คำพังเพย	0	0	0	0	0	0	0	0
08	ศาสนา	0	0	0	0	0	0	0	0
09	ประเพณีวัฒนธรรม	0	0	0	0	0	0	0	0
10	คาถา	0	0	0	0	0	0	0	0
11	หัตถกรรม	0	0	0	0	0	0	0	0
12	การดำรงชีวิต	0	0	0	0	0	0	0	0
13	ดนตรี	1	1	1	1	1	1	1	1
14	การเล่น	0	0	0	0	0	0	1	0
15	การแสดง	0	0	0	0	0	0	1	0

ตารางที่ 3-14 วิเคราะห์ ฟอมอลคอนเท็กซ์ (Formal Context) ของบทความในหมวด 14
การเล่น

		N0001	N0002	N0003	N0004	N0005	N0006	N0007
01	สมุนไพรมะ	0	0	0	0	0	0	0
02	ยารักษาโรค	0	0	0	0	0	0	0
03	ผักและอาหารพื้นบ้าน	0	0	0	0	0	0	0
04	การเลี้ยงเด็ก	0	0	0	0	0	0	0
05	เพลงกล่อมเด็ก	0	0	0	0	0	0	0
06	นิทานพื้นบ้าน	0	0	0	0	0	0	1
07	สุภาอิต คำพังเพย	0	0	0	0	0	0	0
08	ศาสนา	0	0	0	0	0	0	0
09	ประเพณีวัฒนธรรม	0	0	0	0	0	0	1
10	คาถา	0	0	0	0	0	0	1
11	หัตถกรรม	0	0	0	0	0	0	0
12	การดำรงชีวิต	0	0	0	0	0	0	0
13	ดนตรี	0	0	0	0	0	0	1
14	การเล่น	1	1	1	1	1	1	1
15	การแสดง	0	0	0	0	1	0	1

ตารางที่ 3-15 วิเคราะห์ ฟอแมลคอนเท็กซ์ (Formal Context) ของบทความในหมวด 15 การแสดง

		O0001	O0002	O0003	O0004	O0005	O0006
01	สมุนไพร	0	0	0	0	0	0
02	ยารักษาโรค	0	0	0	0	0	0
03	ผักและอาหารพื้นบ้าน	0	0	0	0	0	0
04	การเลี้ยงเด็ก	0	0	0	0	0	0
05	เพลงกล่อมเด็ก	0	0	0	0	0	0
06	นิทานพื้นบ้าน	0	0	1	0	0	0
07	สุภาษิต คำพังเพย	0	0	0	0	0	0
08	ศาสนา	0	0	0	0	0	0
09	ประเพณีวัฒนธรรม	0	0	0	0	0	0
10	คาถา	0	0	0	0	0	0
11	หัตถกรรม	0	0	0	0	0	0
12	การดำรงชีวิต	0	0	0	0	0	0
13	ดนตรี	0	0	1	0	0	0
14	การละเล่น	0	0	1	0	0	0
15	การแสดง	1	1	1	1	1	1

เมื่อจัดรูปแบบข้อมูลเพื่อเตรียมวิเคราะห์ในรูปแบบซีเอสวี (csv) แล้วก็ทำการวิเคราะห์ด้วยโปรแกรม conexp ระหว่าง ออบเจกต์ (Object) จำนวน 15 ออบเจกต์ กับ แอตทริบิวต์ (Attribute) จำนวน 245 แอตทริบิวต์ ดังแสดงไว้ในตารางที่ 3-16

ตารางที่ 3-16 การวิเคราะห์ ฟอแมลคอนเท็กซ์ (Formal Context) ด้วยโปรแกรม conexp

	A0001	A0002	A0003	A0004	A0005	A0006	A0007	A0008	A0009	A0010	A0011	A0012	A0013	A0014
01 สมุนไพร	×	×	×	×	×	×	×	×	×	×	×	×	×	×
02 ยารักษาโรค	×	×	×	×	×	×	×	×	×	×	×	×	×	×
03 ผักและอาหารพื้นบ้าน		×												
04 การเลี้ยงเด็ก							×							
05 เพลงกล่อมเด็ก														
06 นิทานพื้นบ้าน														
07 สุภาษิต คำพังเพย														
08 ศาสนา														
09 ประเพณีวัฒนธรรม														
10 คาถา														
11 หัตถกรรม														
12 การดำรงชีวิต		×												
13 ดนตรี														
14 การละเล่น														
15การแสดง														

จากคอนเซ็ปต์ภายในบริบท (Context) ในตารางที่ 3-1 ถึง 3-16 จะได้คุณสมบัติ

$$X \subseteq G: X \rightarrow X = \{m \in M \mid g|m \text{ สำหรับ } g \in X\}$$

$$Y \subseteq M: Y \rightarrow Y = \{g \in G \mid g|m \text{ สำหรับ } m \in X\}$$

$X = \{01 \text{ สมุนไพร, } 02 \text{ ยารักษาโรค}\}$ $\{02 \text{ ยารักษาโรค, } 10 \text{ คาถา}\}$ $\{01 \text{ สมุนไพร, } 03 \text{ ผักพื้นบ้าน}\}$ $\{04 \text{ การเลี้ยงเด็ก, } 12 \text{ เครื่องมือเครื่องใช้การดำรงชีวิต}\}$ $\{04 \text{ การเลี้ยงเด็ก, } 05 \text{ เพลงกล่อมเด็ก}\}$ $\{06 \text{ นิทานพื้นบ้าน}\}$ $\{07 \text{ สุภาษิต คำพังเพย, } 09 \text{ ประเพณีวัฒนธรรม}\}$

Y:= {A0001,0002,A0003,A0004,A0005,A0006,A0007,A0008,A0009,A0010,A0011, A0012,A0013,A0014,A0015,A0016,A0017,A0018,A0019,A0020,A0021,A0022,A0023,A0024,A0025,A0026,A0027,A0028,A0029,B0001,B0002,B0003,B0004,B0005,B0006,B0007,B0008,B0009,B0010,B0011,C0001,C0002,C0003,C0004,C0005,C0006,C0007,C0008,C0009,C0010,C0011,C0012,C0013,C0014,C0015,C0016,C0017,C0018,C0019,C0020,C0021,C0022,C0023,C0024,C0025,C0026,C0027,C0028,C0029,D0001,D0002,D0003,E0001,E0002,E0003,E0004,E0005,E0006,E0007,E0008,E0009,E0010,E0011,E0012,E0013,E0014,E0015,E0016,E0017,E0018,E0019,E0020,E0021,E0022,E0023,E0024,E0025,E0026,E0027,E0028,E0029,E0030,E0031,E0032,E0033,E0034,E0035,E0036,E0037,E0038,E0039,E0040,F0001,F0002,F0003,F0004,F0005,F0006,G0001,G0002,G0003,G0004,G0005,G0006,G0007,G0008,G0009,G0010,G0011,G0012,H0001,H0002,H0003,H0004,H0005,H0006,I0001,I0002,I0003,I0004,I0005,I0006,I0007,I0008,I0009,I0010,I0011,I0012,J0001,J0002,J0003,J0004,J0005,J0006,J0007,J0008,J0009,J0010,J0011,J0012,J0013,J0014,J0015,K0001,K0002,K0003,K0004,K0005,K0006,K0007,K0008,K0009,K0010,K0011,K0012,K0013,K0014,K0015,K0016,K0017,K0018,K0019,K0020,K0021,K0022,K0023,K0024,K0025,K0026,K0027,K0028,K0029,K0030,L0001,L0002,L0003,L0004,L0005,L0006,L0007,L0008,L0009,L0010,L0011,L0012,L0013,L0014,L0015,L0016,L0017,L0018,L0019,L0020,L0021,L0022,L0023,L0024,L0025,L0026,L0027,L0028,L0029,L0030,L0031,M0001,M0002,M0003,M0004,M0005,M0006,M0007,M0008,N0001,N0002,N0003,N0004,N0005,N0006,N0007,O0001,O0002,O0003,O0004,O0005,O0006 }

3.7.3 แปลงฟอแมอลคอนเท็กซ์ (Formal Context) ให้เป็น คอนเซพแลททิส (Concept lattice) โดยแสดงไว้ในบทที่ 4

3.7.4 ขั้นตอนการสร้าง

3.7.4.1 เริ่มสร้างโหนดแรก แอตทริบิว (Attribute) ที่มี ออบเจกต์ (Object) ร่วมกันมากที่สุดก่อน

3.7.4.2 สร้างโหนดที่มี ออบเจกต์ (Object) ร่วมกันมากที่สุด แต่ไม่มีแอตทริบิวร่วมกันของ Concept ใดๆ

3.7.4.3 สร้างโหนดที่มี ออบเจกต์ (Object) ร่วมกัน แต่ไม่มีแอตทริบิวร่วมกันของ Concept

3.7.4.4 เอาแอตทริบิวตามลักษณะของโหนดแต่ละคู่มา Union กันเพื่อให้ได้โหนดใหม่

3.8 สรุปหมวดหมู่และความสัมพันธ์

3.8.1 สรุปผลหมวดหมู่ และนำหมวดหมู่ที่ได้จากการวิเคราะห์ ไปตรวจสอบโดยผู้ทรงคุณวุฒิ ด้านภูมิปัญญาท้องถิ่นภาคใต้ และสรุปความสัมพันธ์จากคอนเซพแลททิส (Concept Lattice) สำหรับพัฒนาระบบจัดเก็บและสืบค้นภูมิปัญญาท้องถิ่นภาคใต้

3.8.2 นำผลหมวดหมู่และความสัมพันธ์มาสรุปเพื่อเตรียมสำหรับนำไปพัฒนาระบบธนาคาร
 สมองภูมิปัญญาท้องถิ่น

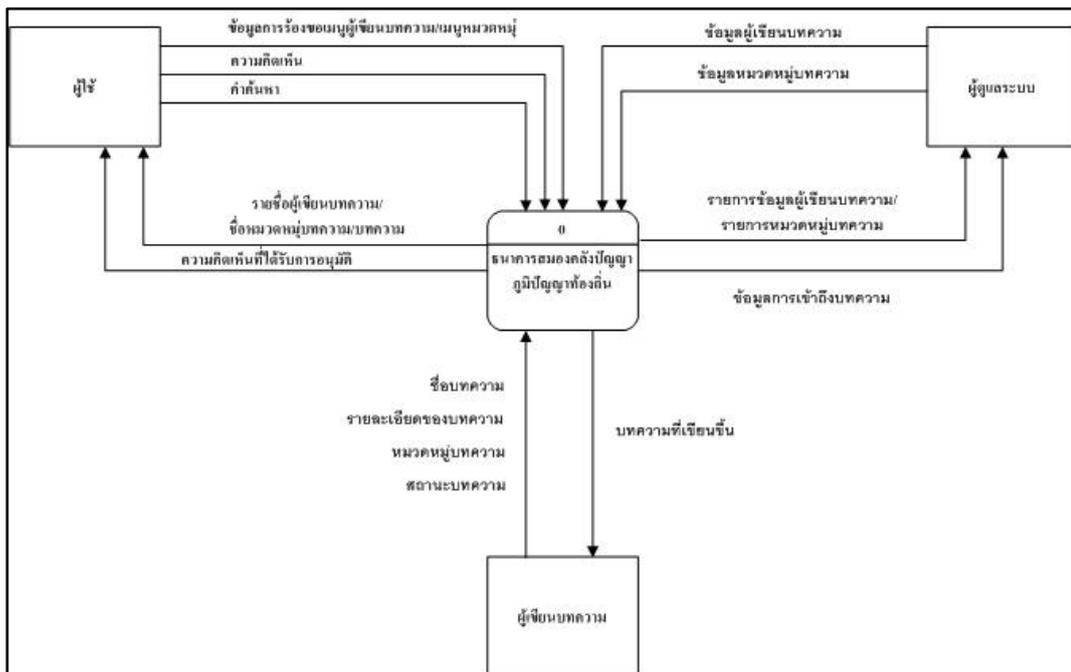
3.8.3 จากคอนเซพแลททิส (Concept Lattice) ออกแบบระบบฐานข้อมูลของระบบ ระบบผัง
 การไหล (DFD) ระบบการจัดเก็บและสืบค้น ประกอบด้วย ระบบการจัดหมวดหมู่บทความ (Cluster
 Analysis) ระบบจัดทำบทความ ระบบพิจารณาความถี่ของคำ (Frequency Transaction) ระบบการ
 สร้างความสัมพันธ์ (Association Rule) และระบบการสืบค้น ออกแบบเชิงแนวคิด (Conceptual
 Design) เชิงตรรกะ (Logical Design) ออกแบบเชิงกายภาพ (Physical Design)

3.9 การออกแบบและพัฒนาระบบ

3.9.1 ออกแบบระบบฐานข้อมูลของระบบจัดเก็บและสืบค้นภูมิปัญญาท้องถิ่นภาคใต้พิจารณา
 ตามความถี่ (Frequency transaction) ของระบบ และหมวดหมู่ที่ได้จากการทำการแบ่งกลุ่มข้อมูล
 (Cluster Analysis) เพื่อหา ความสัมพันธ์ (Association Rule) เพื่อทำ Data Mapping ของระบบ
 เพื่อนำเสนอบทความที่มีความเกี่ยวข้องกัน จากคำค้นเดียวกัน

3.9.2 ออกแบบพัฒนาระบบผังการไหลของระบบโดยพิจารณาตามความถี่ (Frequency
 transaction) ของระบบ และหมวดหมู่ที่ได้จากการทำการแบ่งกลุ่มข้อมูล (Cluster Analysis) เพื่อ
 หา ความสัมพันธ์ (Association Rule) เพื่อทำ Data Mapping ของระบบ เพื่อนำเสนอบทความที่มี
 ความเกี่ยวข้องกัน จากคำค้นเดียวกัน ดังแสดงไว้ในภาพที่ 3-6 ถึง ภาพที่ 3-12

3.9.2.1 Context Diagramme ของระบบ

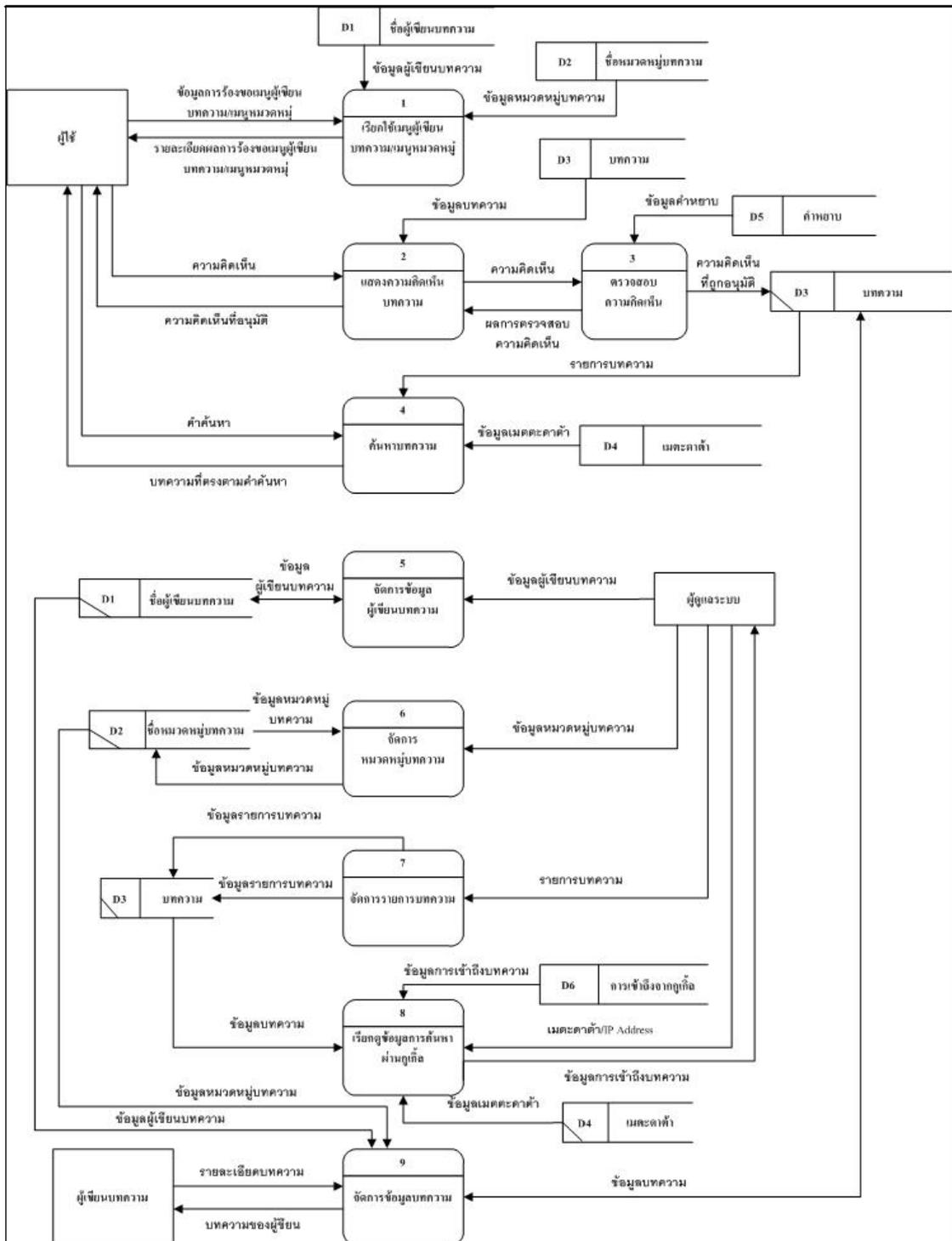


ภาพที่ 3-6 Context Diagramme ของระบบ

จากภาพที่ 3-6 สามารถอธิบายดังนี้

- 1) ผู้ใช้ สามารถค้นหาบทความเพื่อศึกษาบทความที่สนใจ พร้อมแสดงความคิดเห็นต่อบทความที่สนใจ หากผู้ที่มีความประสงค์เป็นผู้เขียนบทความก็สามารถร้องขอผ่าน ผู้ดูแลระบบได้ นอกจากนี้ผู้ใช้อยังสามารถจัดกลุ่มบทความตามหมวดหมู่ที่สนใจ และตามผู้เขียนที่สนใจได้
- 2) ผู้เขียนบทความ สามารถเขียนบทความโดยเลือก เขียนบทความหมวดใดก็ได้ที่มีความถนัดและสนใจ นอกจากนั้นผู้เขียนบทความจะเลือกที่จะไม่เผยแพร่หรือเผยแพร่บทความก็ได้
- 3) ผู้ดูแลระบบ สามารถสร้างผู้เขียนบทความได้ ตามการร้องขอของผู้ใช้ และสามารถลบผู้ใช้ออก หากเผยแพร่บทความที่ไม่เหมาะสม นอกจากนั้นผู้ดูแลระบบสามารถลบบทความที่ไม่เหมาะสมได้ พร้อมทั้งอนุมัติหรือลบความคิดเห็นที่ไม่เหมาะสมที่ผู้ใช้มีต่อบทความได้ ที่สำคัญผู้ดูแลระบบสามารถดูสถิติการอ่านบทความของผู้ใช้ และข้อมูลการถึงบทความจากหมายเลขไอพีแอสแตรสภายนอก

3.9.2.2 DFD Level 1

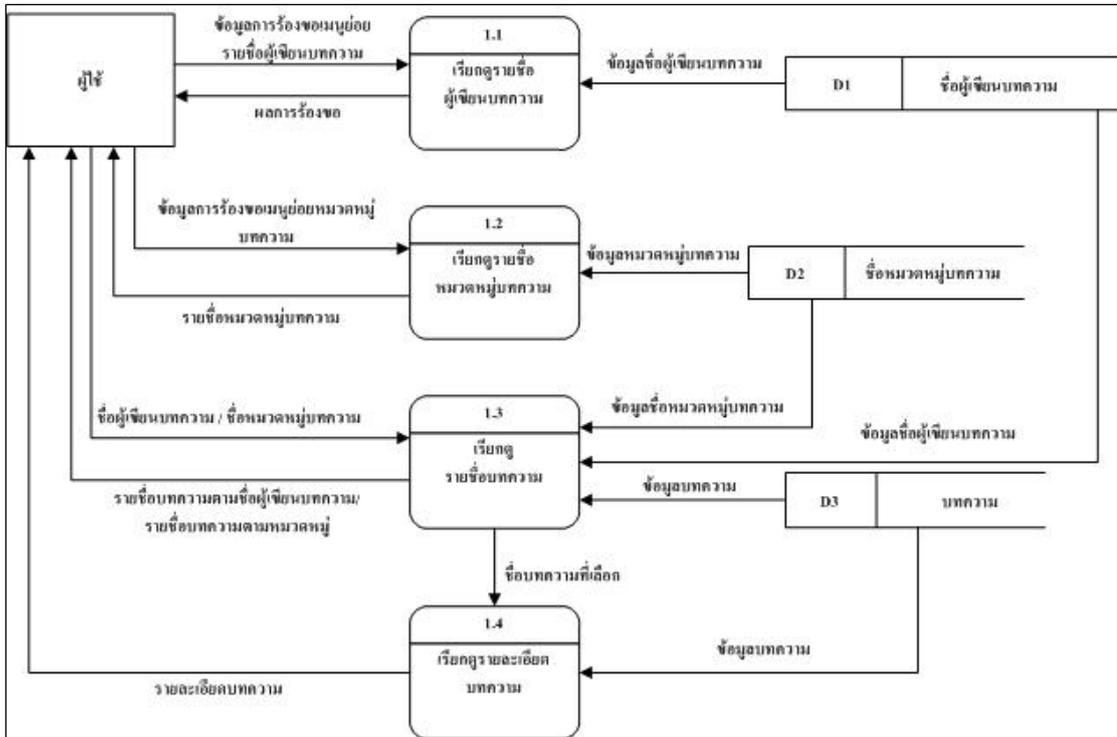


ภาพที่ 3-7 DFD Level 1

จากภาพที่ 3-7 สามารถอธิบายได้ว่า ระบบที่พัฒนาขึ้นประกอบด้วย 9 กระบวนการทำงาน (Process) ดังนี้

- 1) Process 1 ผู้เขียนบทความและจัดหมวดหมู่ เรียกใช้โดยผู้ใช้ที่ต้องการจะเป็นผู้เขียนบทความ มีการจัดเก็บรายชื่อชื่อผู้เขียนบทความ และชื่อหมวดหมู่บทความ
- 2) Process 2 แสดงความคิดเห็นบทความ และ 3 ตรวจสอบความคิดเห็นบทความ เรียกใช้งานโดยผู้ใช้ เพื่อให้ผู้ใช้ได้แสดงความคิดเห็นต่อบทความ และมีการตรวจสอบการแสดงความความคิดเห็นต่อบทความ เพื่อจะอนุมัติหรือปฏิเสธบทความที่จะทำการเผยแพร่
- 3) Process 4 ค้นหาบทความ เรียกใช้งานโดยผู้ใช้ เพื่อค้นหาความที่สนใจตามคำค้นที่ผู้ใช้งานต้องการ คำค้นนั้นจะถูกเปรียบเทียบกับ ข้อมูลเมตาดาต้าของบทความที่จะเป็นตัวแทนของบทความ และแสดงบทความที่เกี่ยวข้องทั้งหมดที่มีการจัดเก็บในข้อมูลเมตาดาต้าแม้จะอยู่ต่างหมวดหมู่
- 4) Process 5 จัดการข้อมูลผู้เขียนบทความ ถูกเรียกใช้โดยผู้ดูแลระบบ เพื่อจัดการข้อมูลผู้เขียนบทความในระบบ
- 5) Process 6 จัดการหมวดหมู่บทความ ถูกเรียกใช้โดยผู้ดูแลระบบ เพื่อจัดการหมวดหมู่ในบทความ ในการเพิ่ม ลบ และแก้ไขหมวดหมู่บทความ
- 6) Process 7 จัดการรายการบทความ ถูกเรียกใช้โดยผู้ดูแลระบบ เพื่อจัดการบทความที่มีการจัดเก็บในระบบ
- 7) Process 8 เรียกดูข้อมูลการค้นหาผ่านกูเกิ้ล ถูกเรียกใช้โดยผู้ดูแลระบบ เพื่อดูข้อมูลการเข้าถึงบทความโดยผู้ใช้งานกูเกิ้ล ซึ่งจะทราบว่าผู้ใช้เข้ามาใช้นั้นเข้าถึงบทความโดยใช้คำค้นอะไรในการค้นผ่านกูเกิ้ลเพื่อมาอ่านบทความ
- 8) Process 9 จัดการข้อมูลบทความ ถูกเรียกใช้โดยผู้เขียนบทความ สามารถเขียนบทความ แก้ไขบทความและเปลี่ยนหมวดหมู่บทความ และกำหนดเมตาดาต้าให้กับบทความ

3.9.2.3 DFD Level 2 Process 1

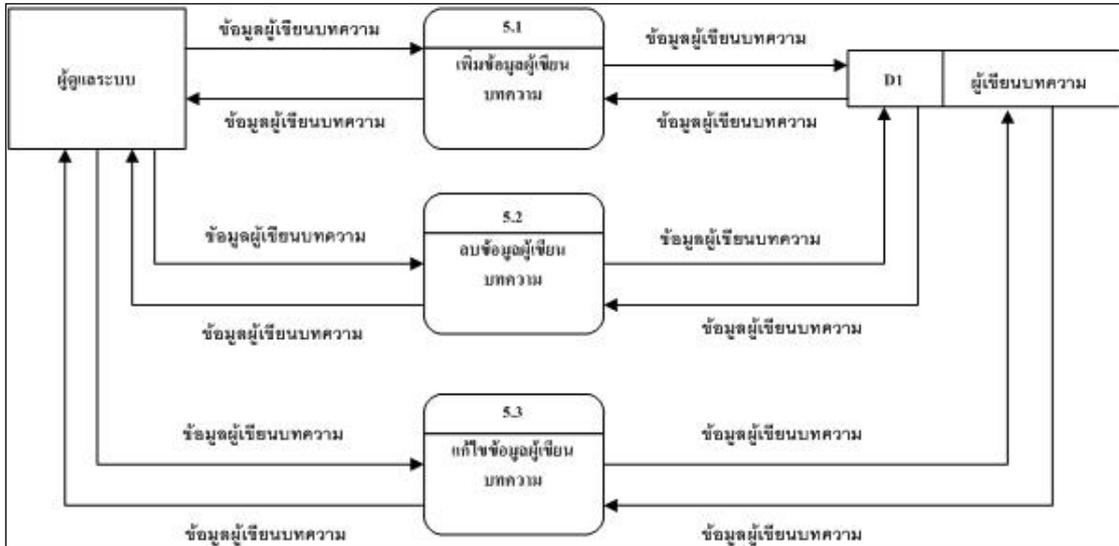


ภาพที่ 3-8 DFD Level Process 1

จากภาพที่ 3-8 DFD Level Process 1 สามารถอธิบายได้ดังนี้

- 1) Process 1.1 เรียกดูรายชื่อผู้เขียนบทความ ถูกเรียกใช้งานโดยผู้ใช้ที่มีความต้องการเป็นผู้เขียนบทความ และมีความต้องการจะเป็นผู้เขียนบทความ มีการจัดเก็บชื่อผู้เขียนบทความ
- 2) Process 1.2 เรียกดูรายชื่อหมวดหมู่บทความ ถูกเรียกใช้โดยผู้ใช้ที่ได้ผ่านอนุมัติเป็นผู้เขียนบทความ มีการจัดเก็บชื่อหมวดหมู่บทความ
- 3) Process 1.3 เรียกดูรายชื่อบทความ ถูกเรียกใช้โดยผู้ใช้ทั่วไปที่มีความสนใจบทความในระบบ มีการจัดเก็บบทความ
- 4) Pprocess 1.4 เรียกดูรายละเอียดบทความ ถูกเรียกใช้โดยผู้ใช้ทั่วไป ที่สนใจค้นหาและอ่านบทความในระบบ มีการจัดเก็บบทความ

3.9.2.4 DFD Level 2 Process 5

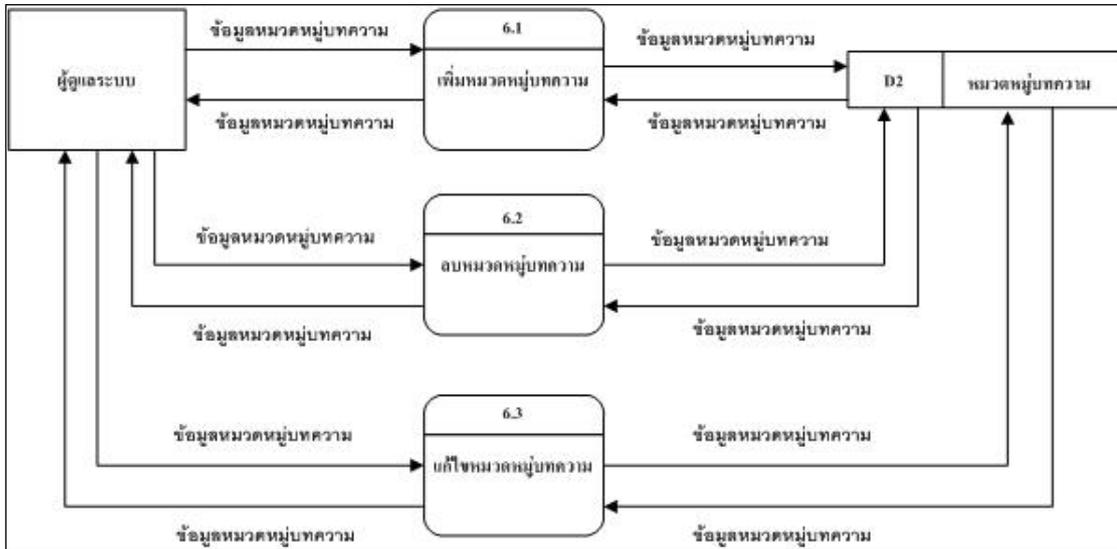


ภาพที่ 3-9 DFD Level Process 5

จากภาพที่ 3-9 DFD Level Process 5 สามารถอธิบายได้ดังนี้

- 1) Process 5.1 เพิ่มข้อมูลผู้เขียนบทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการจัดการข้อมูลผู้เขียนบทความที่ยกฐานะมาจากผู้ใช้ทั่วไป มีการจัดเก็บรายชื่อผู้เขียนบทความ
- 2) Process 5.2 ลบข้อมูลผู้เขียนบทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการลบข้อมูลจากการจัดเก็บรายชื่อผู้เขียนบทความ และลบข้อมูลผู้เขียนบทความที่มีการเขียนบทความที่ไม่เหมาะสม
- 3) Process 5.3 แก้ไขข้อมูลผู้เขียนบทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการจัดการข้อมูลในการแก้ไขข้อมูลจากการจัดเก็บรายชื่อผู้เขียนบทความ และแก้ไขข้อมูลผู้เขียนบทความที่มีการจัดเก็บผิดพลาด

3.9.2.5 DFD Level 2 Process 6

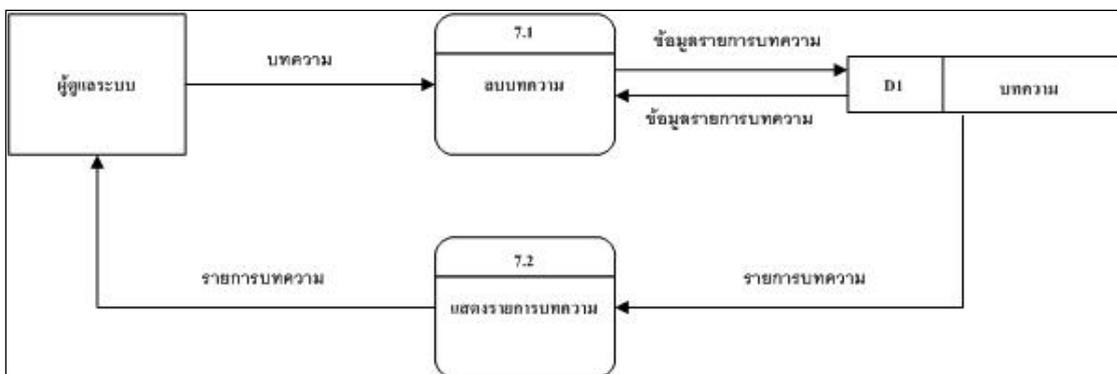


ภาพที่ 3-10 DFD Level Process 6

จากภาพที่ 3-10 DFD Level Process 6 สามารถอธิบายได้ดังนี้

- 1) Process 6.1 เพิ่มหมวดหมู่บทความ มีการจัดเก็บข้อมูลหมวดหมู่บทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการจัดการข้อมูลหมวดหมู่บทความ มีการจัดเก็บข้อมูลหมวดหมู่บทความ
- 2) Process 6.2 ลบหมวดหมู่บทความ มีการจัดเก็บข้อมูลหมวดหมู่บทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการลบข้อมูลหมวดหมู่บทความ
- 3) Process 6.3 แก้ไขหมวดหมู่บทความ มีการจัดเก็บข้อมูลหมวดหมู่บทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการจัดการข้อมูลในการแก้ไขข้อมูลหมวดหมู่บทความ

3.9.2.6 DFD Level 2 Process 7



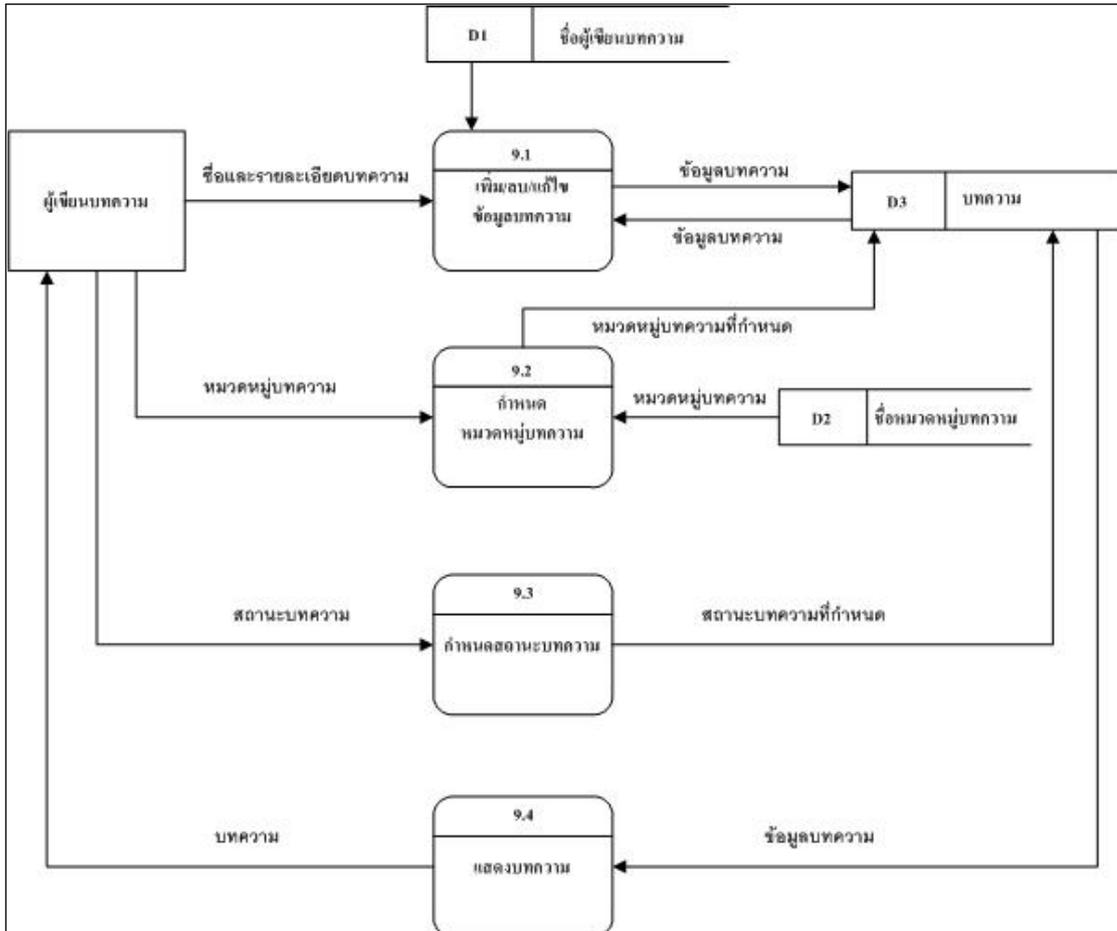
ภาพที่ 3-11 DFD Level Process 7

จากภาพที่ 3-11 DFD Level Process 7 สามารถอธิบายได้ดังนี้

1) Process 7.1 ลบบทความ มีการจัดเก็บข้อมูลบทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการลบบทความในระบบ

2) Process 7.2 แสดงรายการบทความ มีการจัดเก็บข้อมูลบทความ ถูกเรียกใช้โดยผู้ดูแลระบบในการจัดการข้อมูลบทความ

3.9.2.7 DFD Level 2 Process 9



ภาพที่ 3-12 DFD Level Process 9

จากภาพที่ 3-12 DFD Level Process 7 สามารถอธิบายได้ดังนี้

1) Process 9.1 เพิ่ม ลบ แก้ไข ข้อมูลบทความ มีการจัดเก็บข้อมูลบทความ ถูกเรียกใช้โดยผู้เขียนบทความ ในการลบ แก้ไขข้อมูลบทความ

2) Process 9.2 กำหนดหมวดหมู่บทความ มีการจัดเก็บข้อมูลบทความ ถูกเรียกใช้โดยผู้เขียนบทความ ในการกำหนดหมวดหมู่บทความ

3) Process 9.3 กำหนดสถานะบทความ มีการจัดเก็บข้อมูลบทความ ถูกเรียกใช้โดยผู้เขียนบทความ ในการกำหนดสถานะบทความ และกำหนดเมตาดาต้าให้กับบทความ เพื่อใช้ในการสืบค้น

4) Process 9.4 แสดงบทความ มีการจัดเก็บข้อมูลบทความถูกเรียกใช้โดยผู้เขียนบทความ ในการแสดงบทความ

3.9.3 ออกแบบพัฒนาระบบส่วนติดต่อกับผู้ใช้ (Graphic User Interface) และหมวดหมู่ที่ได้จากการทำการแบ่งกลุ่มข้อมูล (Cluster Analysis) และ ความสัมพันธ์ (Association Rule) เพื่อทำ Data Mapping ของระบบ เพื่อนำเสนอบทความที่มีความเกี่ยวข้องกัน จากคำค้นเดียวกัน

3.9.4 พัฒนาระบบจัดเก็บและสืบค้น ระบบการจัดหมวดหมู่บทความ โดยพิจารณาตามความถี่ (Frequency Transaction) ของคำที่ปรากฏเพื่อสร้าง หมวดหมู่เพื่อทำการแบ่งกลุ่มบทความ (Cluster Analysis) พัฒนาระบบความสัมพันธ์ (Association Rule) ของบทความ ด้วยกำหนดเมตาเดตา (Metadata) ให้กับบทความที่มีความเกี่ยวข้องกันของเนื้อหาและคำค้นบางส่วน พัฒนาระบบการสืบค้นเพื่อค้นหาบทความ ที่มีความเกี่ยวข้องกันจากคำค้นเดียวกัน พัฒนาระบบจัดเก็บและสืบค้น

3.10 ทดสอบหาประสิทธิภาพของระบบ

หาประสิทธิภาพของระบบ เป็นการประเมินค่า Recall และ Precision เป็นมาตรวัดการเรียกคืนและความถูกต้องแม่นยำวิธีการประเมินประสิทธิภาพระบบจัดเก็บและสืบค้น ซึ่งเป็นวิธีการที่ใช้ในงานวิจัยของ สถาพร สังข์ศิริ วาทีนี ดวงอ่อนนาม และธนากร สุวรรณวัฒน์

3.10.1 ค่าความระลึก (Recall) สามารถคำนวณจาก สมการที่ 3-5

$$\frac{\text{จำนวนข้อมูลที่ถูกต้องที่ค้นคืนได้}}{\text{จำนวนข้อมูลทั้งหมดที่ค้นคืนออกมาได้}} \times 100 \quad (3-5)$$

ตารางที่ 3-17 การทดสอบประสิทธิภาพ ค่าความระลึก (Recall)

ค่าความระลึก (Recall)	
จำนวนข้อมูลที่ถูกต้องที่ค้นคืนได้	จำนวนข้อมูลทั้งหมดที่ค้นคืนออกมาได้
29	29
11	11
8	29
3	3
40	40
6	6
12	12
4	6
12	12
14	15
28	30
14	31
8	8
4	7
6	6
199	245

$$\frac{199}{245} \times 100$$

เท่ากับ 81%

3.10.2 ค่าความแม่นยำ (Precision) สามารถคำนวณจาก สมการที่ 3-6

$$\frac{\text{จำนวนข้อมูลที่ถูกต้องที่ค้นคืนได้}}{\text{จำนวนข้อมูลที่ถูกต้องทั้งหมดในฐานข้อมูล}} \times 100 \quad (3-6)$$

ตารางที่ 3-18 การทดสอบประสิทธิภาพ ค่าความแม่นยำ (Precision)

ค่าความแม่นยำ (Precision)		
จำนวนข้อมูลที่ถูกต้องที่ค้นคืนได้		จำนวนข้อมูลที่ถูกต้องทั้งหมดในฐานข้อมูล
ตรงกับความต้องการ	ตรงต่อความต้องการบางส่วน	
27	2	29
11	0	12
8	0	10
3	0	3
40	0	40
3	3	6
12	0	12
4	0	5
6	6	12
14	0	15
28	0	28
12	2	12
8	0	8
4	0	5
6	0	6
199		203

$$\frac{199}{203} \times 100$$

เท่ากับ 98%

โดย ใช้คำค้นในการทดสอบการสืบค้นที่เป็นตัวแทนของแต่ละหมวดหมู่ ได้แก่ สมุนไพร ยา ขนมหัศจรรย์ เพลงกล่อมเด็ก ตำนาน สุภาษิต อิสลาม ประเพณี คาถา หัตถกรรม เครื่องมือ เครื่องดนตรี การละเล่น และการแสดง ผลการวิเคราะห์ ได้แสดงไว้ในบทที่ 4 ข้อ 4.5

3.11 สรุปผลประสิทธิภาพของระบบจัดเก็บและสืบค้น

สรุปผลประสิทธิภาพในการสืบค้นของระบบจัดเก็บและสืบค้น จาก การหาค่าความระลึก (Recall) และการหาค่าความแม่นยำ (Precision) เพื่อประเมินผลประสิทธิภาพของระบบจัดเก็บและสืบค้น และได้แสดงไว้ในบทที่ 4 ข้อ 4.5

3.12 สรุปผลการวิจัย

3.12.1 สรุปผลการจัดกลุ่มหมวดหมู่ของบทความจากสถิติการเข้าใช้ โดยใช้เทคนิค K-mean Cluster ตัวแปรเพื่อหาความสัมพันธ์ทาง ตรรกศาสตร์ ค่าความถี่ในการใช้งานของแต่ละตัวแปร

3.12.2 สรุปผลการคาดการณ์ความสนใจบทความภูมิปัญญาท้องถิ่น โดยเทคนิคการค้นหากฎความสัมพันธ์ (Association Rule) โดยพิจารณาจากร้อยละของค่าความเชื่อมั่น และค่าสนับสนุนที่เกิดขึ้น สามารถหาค่าความเชื่อมั่น และค่าสนับสนุน

3.12.3 สรุปผลการวิเคราะห์ฟอแมลคอนเท็กซ์ (Formal Context) จากคอนเซพแลททิซ (Concept Lattice)

3.12.4 สรุปผลหมวดหมู่และกฎความสัมพันธ์ ที่จะใช้พัฒนาระบบจัดเก็บและสืบค้นภูมิปัญญาท้องถิ่นภาคใต้ ได้แก่ ออกแบบระบบฐานข้อมูลของระบบจัดเก็บและสืบค้นภูมิปัญญาท้องถิ่นภาคใต้ ออกแบบพัฒนาระบบผังการไหลของระบบ และออกแบบพัฒนาระบบส่วนติดต่อกับผู้ใช้ (Graphic User Interface)

3.12.5 สรุปผลนำผลการวิเคราะห์ไปพัฒนาระบบจัดเก็บและสืบค้น

3.12.6 สรุปผลประสิทธิภาพในการสืบค้นของระบบจัดเก็บและสืบค้น จาก การหาค่าความระลึก (Recall) และการหาค่าความแม่นยำ (Precision) เพื่อประเมินผลประสิทธิภาพของระบบจัดเก็บและสืบค้น

3.12.7 สรุปการพัฒนาบบจัดเก็บและสืบค้นภูมิปัญญาท้องถิ่นภาคใต้ ตั้งชื่อระบบว่า ธนาคารสมองภูมิปัญญาท้องถิ่น