



การค้นคืนสารสนเทศตามการจัดหมวดหมู่ระบบทศนิยมดิวอี้แบบสหความสัมพันธ์

วิไลพร เลิศมหาเกียรติ* และ อนิราช มิ่งขวัญ**

บทคัดย่อ

การพัฒนาอย่างรวดเร็วของเทคโนโลยีมีส่งผลให้ผู้ใช้ต้องเผชิญกับข้อมูลปริมาณมหาศาลทั้งที่ตรงและไม่ตรงความต้องการ ในงานวิจัยนี้มุ่งที่นำเทคโนโลยีมาผสมผสานกับกรอบความรู้ในการวิเคราะห์สารสนเทศของห้องสมุดโดยนำเทคโนโลยีมาใช้ในการสกัดข้อมูลสารสนเทศจากหนังสือมาจัดหมวดหมู่ด้วยเทคนิคการวิเคราะห์หมวดหมู่ระบบทศนิยมดิวอี้แบบสหความสัมพันธ์ (Multiple DDC Relational Classification) โดยนำกรอบความรู้ระบบทศนิยมดิวอี้มาใช้เป็นเครื่องมือในการแสดงความสัมพันธ์ของสัดส่วนเนื้อหาของวัสดุสารสนเทศและใช้พีชคณิตในการเปรียบเทียบผลในการค้นคืน เพื่อนำไปเสนอผลลัพธ์ให้กับผู้ใช้ด้วยกราฟิก ซึ่งจะส่งผลให้เพิ่มความแม่นยำและแยกแยะผลของการค้นคืนได้ชัดเจนขึ้น

คำสำคัญ: การค้นคืนสารสนเทศ การวิเคราะห์หมวดหมู่ระบบทศนิยมดิวอี้ การจัดหมู่ดิวอี้แบบสหความสัมพันธ์

1. บทนำ

ในยุคเทคโนโลยีสารสนเทศวิทยาการด้านคอมพิวเตอร์ได้ส่งผลกระทบต่อเปลี่ยนแปลงในวิชาชีพและวิถีทางการดำเนินงานในองค์กร โดยเฉพาะอย่างยิ่งองค์กรทางการศึกษาและข้อมูลข่าวสารจะได้รับผลกระทบมาก เนื่องจากเทคโนโลยีได้เพิ่มประสิทธิภาพในการผลิตสื่อในหลากหลายรูปแบบในปริมาณมหาศาล และส่งผ่านความรู้และข่าวสารได้อย่างรวดเร็วในหลากหลายช่องทาง ส่งผลให้เกิดแนวคิดและการดำเนินงานด้านสารสนเทศที่ปรับเปลี่ยนไป ประการแรกเกิดจากเนื้อหาสารสนเทศมีความหลากหลาย และเป็นสหสาขาวิชามากขึ้นได้ถูกบันทึกในรูปแบบที่แตกต่างกัน

มิได้จำกัดเพียงรูปแบบหนังสือเท่านั้น หากรวมถึงวัสดุสารสนเทศที่ผลิตในสมัยปัจจุบันทุกรูปแบบเป็นความรู้ซึ่งบันทึกลงไว้เพื่อใช้ในการตัดสินใจ ประการต่อมาการค้นคืนสารสนเทศมิได้มีขอบเขตเพียงการมองหาสิ่งใดสิ่งหนึ่งว่าอยู่ที่ใดเท่านั้น แต่รวมถึงการประมวลความรู้ที่ได้บันทึกไว้เนื่องจากกับผู้ใช้มีความต้องการสารสนเทศในระดับที่ลึกถึงตัวเนื้อหาและข้อมูลที่ถูกประมวลไว้เพื่อการตัดสินใจ ประการสุดท้ายระบบการค้นคืนสารสนเทศ (Information retrieval system) ซึ่งแต่เดิมดำเนินการโดยมนุษย์ผู้มีความเชี่ยวชาญในการกำหนดหมวดหมู่เนื้อหาสารสนเทศปรับเปลี่ยนเป็นการนำเทคโนโลยีมาใช้เพื่อความรวดเร็วและความแม่นยำในการค้นคืน อย่างไรก็ตามมนุษย์ได้รับการยอมรับว่ามีความสามารถในการจัดทำดัชนีได้ดีกว่าเทคโนโลยี ในขณะที่เทคโนโลยีมีประสิทธิภาพด้านการประเมินผลที่รวดเร็วกว่าการดำเนินงานเกี่ยวข้องกับความรู้และสารสนเทศทั้งในห้องสมุดและอินเทอร์เน็ต จึงมีความจำเป็นที่จะต้องปรับเปลี่ยนการทำงานให้ทันกับปริมาณสารสนเทศที่เพิ่มขึ้นอย่างมหาศาล และส่งถ่ายสารสนเทศเหล่านั้นไปยังผู้ใช้ที่มีความต้องการสารสนเทศในระดับลึกกว่าเดิม ตลอดจนแก้ปัญหาที่ก่อให้เกิดความรำคาญต่อผู้ใช้ในด้านขั้นตอนในการเข้าถึงสารสนเทศและค้นคืนวัสดุสารสนเทศที่ดีที่สุดเพื่อนำเสนอผู้ใช้งานค้นหาสารสนเทศที่เกี่ยวข้องเพื่อเสนอแนะต่อผู้ใช้ เพื่อให้สามารถค้นหาข้อมูลต่อไปได้ [1]

งานวิจัยนี้มุ่งที่จะนำเสนอแนวทางในการค้นคืนสารสนเทศในรูปแบบการเปรียบเทียบความคล้ายคลึงของสัดส่วนความสัมพันธ์เนื้อหาในวัสดุสารสนเทศกับค่าค้นของผู้ใช้ภายใต้กรอบความรู้ตามระบบมาตรฐานทศนิยมดิวอี้ (DDC) ซึ่งมีการเรียงเป็นลำดับชั้น มีการใช้สัญลักษณ์ระบบตัวเลขที่ง่ายสำหรับทั้งบรรณารักษ์ และผู้อ่านที่จะเรียนรู้วิธีใช้ระบบ [2] -

* ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

** ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

[5] โดยนำเทคโนโลยีในการสกัดข้อมูลสารสนเทศจากหนังสือมาจัดหมวดหมู่ด้วยเทคนิคการวิเคราะห์หมวดหมู่แบบสหความสัมพันธ์ (Multiple DDC Relational Classification) ซึ่งจะทำให้สามารถแสดงให้เห็นสัดส่วนความสัมพันธ์ที่ชัดเจนตลอดจนสร้างฐานข้อมูลสัดส่วนความสัมพันธ์ของเนื้อหาของวัสดุสารสนเทศ เพื่อพัฒนารูปแบบการค้นคืนสารสนเทศที่มีประสิทธิภาพด้านความแม่นยำในการค้นคืนสูงขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การค้นคืนสารสนเทศ

การค้นคืนสารสนเทศ (Information retrieval) หมายถึง กระบวนการ ที่เกี่ยวข้องกับการค้นหาและนำสารสนเทศที่ตรงตามความต้องการ ส่งให้แก่ผู้ใช้อย่างรวดเร็ว ระบบค้นคืนสารสนเทศจึงมีความเกี่ยวข้องกับการจัดเก็บข้อมูล การเข้าถึงข้อมูล และรูปแบบการแทนข้อมูล โดยมีจุดมุ่งหมายคือการค้นคืนเอกสารที่ตรงตามความต้องการของผู้ใช้มากที่สุด

2.1.1 ระบบการค้นคืนสารสนเทศในห้องสมุด ดำเนินงานโดยผู้เชี่ยวชาญซึ่งเน้นวิธีการวิเคราะห์จากเนื้อหา และรูปแบบของหนังสือเป็นหลักแล้วกำหนดสัญลักษณ์ให้หนังสือแต่ละเล่ม ให้อยู่ในกลุ่มหรือหมวดหมู่ใดเพียงหมวดเดียว จากนั้นจัดทำเครื่องมือช่วยค้นในรูปแบบบัตรรายการและดัชนี เพื่อใช้บอกตำแหน่งที่ตั้งของหนังสือในห้องสมุดให้กับผู้ใช้ ซึ่งการดำเนินงานทำด้วยระบบมือทั้งหมดเป็นงานที่ยุ่งยากและต้องใช้เวลา ตลอดจนถึงอาศัยผู้เชี่ยวชาญ ทำให้มีค่าใช้จ่ายสูงและเกิดความล่าช้า

2.1.2 ระบบการค้นคืนสารสนเทศในอินเทอร์เน็ตที่นิยมใช้ ในปัจจุบันคือระบบเครื่องมือสืบค้น (Search Engine System) ซึ่งเป็นกระบวนการในการค้นหาสารสนเทศที่ต้องการจากกลุ่มของเอกสาร (Collections) โดยจะรับคำขอ (Query) ของผู้ใช้ แล้วนำไปค้นหาจากกลุ่มและค้นคืนเป็นรายการของเอกสารที่สัมพันธ์กับคำขอ ซึ่งการทำงานของระบบการค้นคืนสารสนเทศจะประกอบด้วยรวบรวมข้อมูล การสร้างตัวแทนของเอกสาร การเก็บข้อมูล และวิธีการค้นหาสารสนเทศ ซึ่งให้ความสำคัญกับการทำดัชนีคำค้น ด้วยเทคนิคต่าง ๆ ของ Information Retrieval ทำให้เพิ่มประสิทธิภาพในการค้นหาที่รวดเร็วมากขึ้น แต่การจัดทำดัชนีคำค้นโดยเครื่องกล ส่งผลให้ผู้ใช้ต้องเผชิญกับข้อมูลปริมาณมหาศาลที่ตรงความต้องการและไม่ตรงความต้องการ

2.2 รูปแบบการแสดงผลการค้นคืนสารสนเทศ

รูปแบบการแสดงผลการค้นคืนสารสนเทศทั้งในส่วนระบบอัตโนมัติในห้องสมุดและระบบเครื่องมือสืบค้นบนอินเทอร์เน็ต ส่วนใหญ่แสดงผลในรูปแบบรายการสารสนเทศที่มีความเกี่ยวข้องกับคำค้น หรือคำขอ (queries) ของผู้ใช้ตามลำดับตัวอักษร หรือตามปีที่พิมพ์ หรือตามความนิยมและความถี่ในการถูกเรียกใช้ การนำกราฟิกมาใช้ในการแสดงผลจะเป็นรูปแบบใหม่ของเครื่องมือสืบค้น ที่เร้าความสนใจด้วยสีสันสวยงามตื่นตาตื่นใจและมีความชัดเจนของกลุ่มข้อมูลที่ต้องการแสดงด้วยระบบทศนิยมดิวอี้ ซึ่งผู้ใช้ห้องสมุดมีความคุ้นเคย ตลอดจนถึงการแนะนำคำค้นในระบบ เพื่อให้ผู้ใช้ได้สารสนเทศที่มีคุณค่าตรงกับความต้องการให้มากขึ้น [6]

2.3 งานวิจัยที่เกี่ยวข้อง

งานวิจัยด้านการค้นคืนของเครื่องมือสืบค้นที่เกี่ยวข้องกับการแสดงผลส่วนใหญ่มุ่งเน้นศึกษาองค์ประกอบแต่ละส่วนที่จำเป็นและช่วยผู้ค้นคืนให้ได้ข้อมูลที่ตรงความต้องการโดยเร็วที่สุด [7] การกำหนดขนาดหัวข้อที่ต่างกันในเรื่องของรายการผลลัพธ์ [8] สำหรับงานวิจัยด้านการใช้กราฟิกในการแสดงผลการสืบค้นได้มีการใช้เทคนิค 2D และ Flash เพื่อให้การแสดงผลที่ชัดเจน สวยงามและดูน่าสนใจ นอกจากนี้ยังมีการจัดกลุ่มสารสนเทศแล้วนำมาแสดงผลด้วยเทคนิค spiral ซึ่งแสดงเอกสารที่มีความเกี่ยวข้องกับคำค้นเรียงเป็นวงกลมตามระดับความคล้ายคลึง [9] ตลอดจนถึงพัฒนาเทคนิค Cluster Bulls-Eye และ RankSpiral เพื่อใช้ในการแสดงผลเอกสารที่เกี่ยวข้องจากเครื่องมือสืบค้นต่าง ๆ โดยเอกสารที่มีความเกี่ยวข้องสูงมากจะอยู่ใกล้จุดศูนย์กลาง [10]

ในส่วนงานวิจัยด้านการพัฒนากระบวนการค้นคืนสารสนเทศนั้นมีการนำเทคนิคการสร้างกลุ่มคำศัพท์ที่สอดคล้องกับคำศัพท์สัมพันธ์ทางด้านวิศวกรรม (Ei) มาใช้ในการจำแนกเว็บไซต์ตามกลุ่มสาขาทางด้านวิศวกรรมศาสตร์ [11] และมีการนำกรอบความรู้ตามระบบมาตรฐานทศนิยมดิวอี้ (DDC) และระบบมาตรฐานหอสมุดรัฐสภาอเมริกัน (LLC) มาทดสอบจัดหมวดหมู่ให้เว็บไซต์ [12], [13]

งานวิจัยที่กล่าวมาส่วนใหญ่เป็นการศึกษาการค้นคืนบนอินเทอร์เน็ตซึ่งแสดงผลเน้นการนำเสนอเอกสารที่มีความเกี่ยวข้องตามลำดับ (Ranking) ของการใช้หรือ keywords ในขณะที่งานวิจัยด้านการจัดเอกสารตามระบบมาตรฐานทศนิยมดิวอี้ (DDC) มีการศึกษาเฉพาะเรื่องหรือสาขา และการ



จัดตามหมวดหมู่หลักเท่านั้น ซึ่งในงานวิจัยนี้นำกรอบแผนการจัดหมู่หนังสือระบบคณิตวิธีที่ใช้สัญลักษณ์เป็นระบบตัวเลข ซึ่งเป็นที่นิยมในห้องสมุดมาใช้ในการควบคุมคำศัพท์และคำสำคัญที่ปรากฏในสารสนเทศของหนังสือและวัสดุห้องสมุด จากนั้นจึงนำมาหาสัดส่วนความสัมพันธ์ของเนื้อหาสารสนเทศของหนังสือตามหมวดหมู่ระบบ DDC และแสดงผลการค้นคืนด้วยกราฟิก ซึ่งจะกล่าวถึงในส่วนที่ 3 วิธีการดำเนินการวิจัย และอธิบายส่วนประกอบของระบบงานในตอนท้ายที่ 4 ต่อไป

3. วิธีการค้นคืนสารสนเทศรูปแบบสหความสัมพันธ์

จากความต้องการพัฒนาการค้นคืนสารสนเทศในรูปแบบสหความสัมพันธ์ (Multiple Relational) ซึ่งจะสามารถแสดงเนื้อหาของสารสนเทศได้หลายหมวดหมู่ให้กับผู้ใช้ ในงานวิจัยนี้จึงแบ่งงานออกเป็น 2 ส่วนคือ ส่วนการวิเคราะห์วัสดุสารสนเทศเพื่อการจัดเก็บความสัมพันธ์ของเนื้อหาตามกรอบความรู้ของการจัดหมวดหมู่ระบบคณิตวิธี และส่วนของการวิเคราะห์คำค้นเพื่อนำมาเปรียบเทียบในการค้นคืน

3.1 ในส่วนการวิเคราะห์วัสดุสารสนเทศเพื่อการจัดเก็บความสัมพันธ์ของเนื้อหาตามกรอบความรู้ของการจัดหมวดหมู่ระบบคณิตวิธี เป็นขั้นตอนการหาสัดส่วนของเนื้อหาของหนังสือว่าอยู่ในหมวดหมู่ DDC ไດบ้าง ในปริมาณเท่าใด ซึ่งการจะหาสัดส่วนของหมวดหมู่ของเนื้อหาสารสนเทศได้นั้นจะต้องทราบว่าหนังสือแต่ละเล่มนั้นมีเนื้อหาเกี่ยวกับเรื่องใดบ้าง การนำหนังสือทั้งเล่มมาใช้ในการศึกษาจะทำให้ต้องวิเคราะห์คำสำคัญจำนวนมากเกินความจำเป็น ในงานวิจัยนี้ผู้วิจัยจึงนำข้อมูลจากสารบัญและดัชนีท้ายเล่มของหนังสือมาสกัดหาคำสำคัญของหนังสือตามกระบวนการของ Information Retrieval เนื่องจากสารบัญของหนังสือประกอบด้วยหัวเรื่องใหญ่ที่บ่งบอกว่าหนังสือนั้นครอบคลุมในเรื่องใดบ้าง ในขณะที่ดัชนีคำค้นท้ายเล่มหนังสือประกอบด้วยคำสำคัญที่ปรากฏในเนื้อหาของหนังสือ

กรอบความรู้ที่นำมาใช้ในการวิเคราะห์หมวดหมู่สารสนเทศเพื่อให้สามารถเห็นความสัมพันธ์เป็นลำดับ ผู้วิจัยจึงเลือกระบบมาตรฐานคณิตวิธี (DDC) ซึ่งเป็นระบบที่มีการจัดกรอบองค์ความรู้ที่เป็นลำดับชั้นและกำหนดสัญลักษณ์ที่เป็นลำดับชั้นอย่างชัดเจน นอกจากนี้ระบบนี้ยังมีดัชนีสัมพันธ์ (Relative Index) เพื่อช่วยในการเข้าถึงตารางการจัดหมู่ของระบบได้อย่างรวดเร็ว ดังนั้นการหาสัดส่วนสหความสัมพันธ์ของเนื้อหาของหนังสือด้วยระบบ DDC จะทำให้เห็นกรอบ

เนื้อหาของหนังสือได้อย่างชัดเจนว่ามีความสัมพันธ์กับหมวดใดหมู่ใดในสัดส่วนเท่าใด

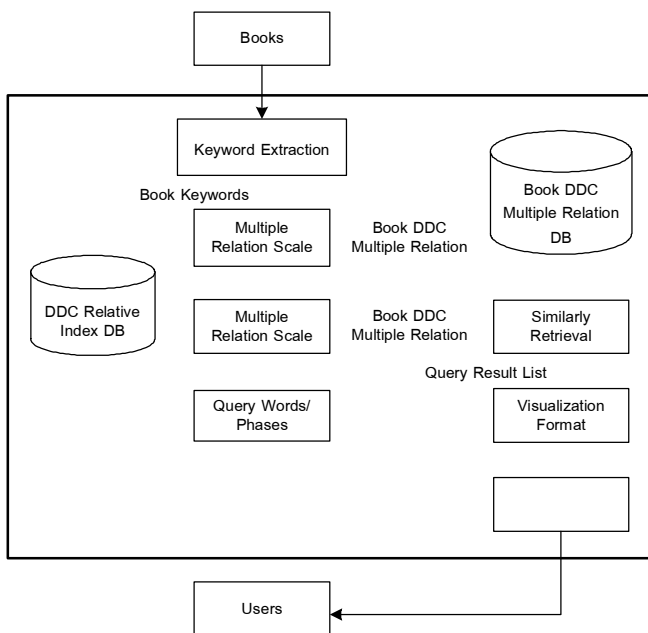
3.2 ในส่วนของการวิเคราะห์คำค้นเพื่อนำมาเปรียบเทียบในการค้นคืน เป็นขั้นตอนของการหาสัดส่วนของเนื้อหาของคำค้น (Query) ว่าอยู่ในหมวดหมู่ DDC ไດ ซึ่งต้องนำคำค้นของผู้ใช้ทั้งในรูปแบบคำวลี มาสกัดหาคำสำคัญของคำค้นของผู้ใช้แล้วนำไปวิเคราะห์หาความสัมพันธ์และสัดส่วนของหมวดหมู่ระบบคณิตวิธี

การนำเสนอรายการผลลัพธ์ของการค้นคืนสารสนเทศนั้นใช้วิธีการเสนอตามลำดับความคล้ายคลึงของสัดส่วนสหความสัมพันธ์ของเนื้อหาของหนังสือ (Books DDC Multiple Relation) กับสัดส่วนของเนื้อหาของคำค้น (Books DDC Multiple Relation) เนื่องจากรายละเอียดในหมวดหมู่ DDC ของสัดส่วนความสัมพันธ์เนื้อหาของหนังสือมีมากกว่าสัดส่วนความสัมพันธ์ของคำค้น โดยในที่นี้ผู้วิจัยใช้เทคนิคพีชชีมาช่วยในการแจกแจงและเปรียบเทียบระดับความคล้ายคลึงของรูปแบบสัดส่วนความสัมพันธ์ระหว่างหนังสือและคำค้นของผู้ใช้ รูปแบบการแสดงผลในการค้นคืนเป็นอีกส่วนที่มีความสำคัญ เพราะเป็นส่วนที่ติดต่อกับผู้ใช้และต้องสามารถนำเสนอในรูปแบบของรายการหนังสือที่มีความคล้ายคลึงกับสัดส่วนของคำค้นตามลำดับค่าความคล้ายคลึงของหมวดหมู่ คณิตวิธี ซึ่งต้องสามารถแสดงได้หลายหมวดหมู่ ผู้วิจัยจึงใช้กราฟิกในการนำเสนอซึ่งสามารถแสดงความสัมพันธ์ในหลายรูปแบบได้

4. การออกแบบผังระบบงาน

จากการวิเคราะห์ค้นคืนสารสนเทศรูปแบบสหความสัมพันธ์ ผู้วิจัยได้นำมาออกแบบผังระบบงานโดยแบ่งส่วนประกอบออกเป็น 2 ส่วน คือ ส่วนเตรียมฐานข้อมูล และส่วนการค้นหาเอกสารของผู้ใช้ ในส่วนเตรียมฐานข้อมูลผู้วิจัยนำหนังสือมาเป็นกลุ่มตัวอย่างในการศึกษาความสัมพันธ์ของเนื้อหาสารสนเทศ โดยเก็บข้อมูลจากหน้าสารบัญและดัชนีท้ายเล่ม เนื่องจากส่วนสารบัญประกอบด้วยกลุ่มคำที่เป็นประเด็นสำคัญของเนื้อหา ในขณะที่ดัชนีท้ายเล่มนั้นจะระบุคำสำคัญที่ปรากฏในเนื้อเรื่องของหนังสือ การเก็บข้อมูลจากทั้งสองส่วนจึงเป็นตัวแทนที่ดีที่ครอบคลุมเนื้อหาทั้งหมด จากนั้นเมื่อนำมาผ่านกระบวนการด้าน Information Retrieval เพื่อจัดทำดัชนีและวิเคราะห์หมวดหมู่เพื่อหาสัดส่วนความ

สัมพันธ์ของเนื้อหาสารสนเทศภายใต้กรอบความรู้ระบบมาตรฐานทศนิยมดิวอี้จัดเก็บในฐานข้อมูล ในส่วนของการค้นคืนเป็นนำคำค้น (queries) ของผู้เข้ามาเพื่อหาสัดส่วนความสัมพันธ์ของหมวดหมู่เนื้อหาภายใต้กรอบความรู้ระบบมาตรฐานทศนิยมดิวอี้ แล้วนำไปเปรียบเทียบความคล้ายคลึงของสัดส่วนความสัมพันธ์ของเนื้อหาที่จัดเก็บไว้ในฐานข้อมูล และค้นคืนให้กับผู้ใช้ตามลำดับความเหมือน (similarity) พร้อมนำเสนอหนังสือที่มีเนื้อหาที่มีความเกี่ยวข้อง โดยมีกระบวนการดำเนินงานตามภาพผังระบบงานดังภาพที่ 1



ภาพที่ 1 ผังระบบงานการค้นคืนรูปแบบสหความสัมพันธ์

4.1.1 การนำข้อมูลหน้าสารบัญและหน้าดัชนีของหนังสือ ซึ่งประกอบด้วยกลุ่มคำและคำสำคัญ มาสกัดและแปลงให้อยู่ในรูปแบบที่ต้องการ (representation) โดยการตัดคำที่ไม่จำเป็น (stop words) การทำให้อยู่ในรูปของรากศัพท์ (stemming) และการหาคำเหมือน (normalization) การให้นำหนักของคำสำคัญ เพื่อให้ได้คำสำคัญของหนังสือ (Book Keywords)

4.1.2 นำคำสำคัญของหนังสือมาวิเคราะห์หมวดหมู่แบบสหความสัมพันธ์ โดยใช้คู่มือการจัดหมวดหมู่ระบบทศนิยมดิวอี้มาเป็นเครื่องมือ เพื่อหาสัดส่วนของเนื้อหาสารสนเทศของหนังสือภายใต้กรอบความรู้ระบบมาตรฐานทศนิยมดิวอี้ จัดเก็บไว้ในฐานข้อมูลสหความสัมพันธ์ (Books

DDC Multiple Relational) เพื่อใช้ในการเปรียบเทียบกับ query ของผู้ด้วยเทคนิคฟัซซี่

4.1.3 ผู้ใช้ใส่คำค้น (queries) ที่ต้องการทั้งในรูปแบบคำ(words) หรือวลี (phases) จากนั้นนำไปหาคำสำคัญ (keywords) เพื่อให้ให้นำหนักของคำ

4.1.4 นำคำสำคัญของคำค้นมาวิเคราะห์หมวดหมู่แบบสหความสัมพันธ์ด้วยดัชนีสัมพันธ์ของระบบจัดหมู่ทศนิยมดิวอี้เพื่อหาสัดส่วนของเนื้อหาสารสนเทศของคำค้นภายใต้กรอบความรู้ระบบมาตรฐานทศนิยมดิวอี้

4.1.5 นำฟัซซี่มาใช้ในการเปรียบเทียบความคล้ายคลึงสัดส่วนความสัมพันธ์ตามมาตรฐานระบบทศนิยมดิวอี้ของคำค้นกับสัดส่วนความสัมพันธ์ตามมาตรฐานระบบทศนิยมดิวอี้ของหนังสือในฐานข้อมูล เพื่อนำเสนอรายการผลลัพธ์ตามลำดับสัดส่วนความสัมพันธ์

4.1.6 นำรายการผลลัพธ์ตามลำดับสัดส่วนความสัมพันธ์ตามมาตรฐานระบบทศนิยมดิวอี้ไปนำเสนอต่อผู้ใช้ในรูปแบบกราฟิก

5. ผลการดำเนินงาน

จากการออกแบบเพื่อพัฒนารูปแบบการค้นคืนสารสนเทศแบบสหความสัมพันธ์ ผู้วิจัยทำการสำรวจหน้าจอแสดงผลของเครื่องมือสืบค้นพบว่ามียังประกอบหลักที่พบในเครื่องมือสืบค้นทุกตัวคือหัวข้อเรื่อง (Title) เนื้อหาโดยสรุป (Summary) และที่อยู่ของแหล่งข้อมูล (URL) ส่วนประกอบอื่น ๆ เป็นองค์ประกอบเสริมซึ่งจะช่วยเสริมให้เครื่องมือสืบค้นนั้นมีลักษณะเด่น สวยงาม และใช้งานได้สะดวกรวดเร็วขึ้น โดยเฉพาะการนำกราฟิกมาใช้ในการแสดงผลร่วมกับการจัดกลุ่มสารสนเทศตามกรอบความรู้ของระบบทศนิยมดิวอี้ จะทำให้เพิ่มประสิทธิภาพในการค้นคืนและช่วยแนะแนวทางในการค้นคืนให้กับผู้ใช้ [6] โดยแสดงผลหน้าจอผู้ใช้ใน 3 ส่วนหลัก คือ ส่วนแรกตำแหน่งซ้ายบนของหน้าจอแสดงรายการเอกสารหลัก 5 รายการตามลำดับความสัมพันธ์กับคำค้น ส่วนที่สองตำแหน่งขวาบนแสดงรายการคำค้นอื่นในหมวดหมู่ระบบ DDC ที่มีความเกี่ยวข้องตามลำดับความสัมพันธ์กับคำค้น ในส่วนที่ 3 ด้านล่างของหน้าจอแสดงผลหมวดหมู่ DDC 10 หมวดหลัก ดังภาพที่ 2



ภาพที่ 2 รูปแบบหน้าจอการแสดงผลการค้นคืนแบบสหความสัมพันธ์

6. บทสรุป

การวิเคราะห์หมวดหมู่หนังสือของผู้เชี่ยวชาญด้านห้องสมุดนิยมนิเคราะห์จากเนื้อหา และรูปแบบของวัสดุสารสนเทศเป็นหลัก จากนั้นจึงกำหนดสัญลักษณ์ให้หนังสือเล่มนั้นๆ ให้อยู่ในกลุ่มหรือหมวดหมู่ใดเพียงหมวดเดียวเท่านั้น เนื่องจากข้อจำกัดของรูปแบบวัสดุสารสนเทศและการระบุตำแหน่งที่ตั้งหนังสือในห้องสมุด ทำให้เนื้อหาสารสนเทศอื่นที่มีในวัสดุสารสนเทศไม่ได้รับการวิเคราะห์หมวดหมู่และการนำเสนอผลการสืบค้นยังเป็นการแสดงรายการวัสดุสารสนเทศตามรายการเฉพาะที่มีสัญลักษณ์เท่านั้น ผู้วิจัยจึงนำเทคโนโลยีมาใช้ในการวิเคราะห์สารสนเทศในห้องสมุดด้วยเทคนิคการวิเคราะห์หมวดหมู่แบบสหความสัมพันธ์ (Multiple DDC Relational Classification) จะทำให้สารสนเทศที่เคยมองไม่เห็นหรือไม่สามารถแสดงผลได้ ให้สามารถนำกลับมาเป็นสารสนเทศและนำเสนอต่อผู้ใช้ในรูปแบบความสัมพันธ์ที่ชัดเจนขึ้นว่าหนังสือหรือเอกสารหนึ่งๆ มีความสัมพันธ์กับเรื่องใดบ้างในปริมาณเท่าใดและมีรูปแบบความสัมพันธ์อย่างไร รวมทั้งสามารถนำเสนอเอกสารที่มีรูปแบบมิติความสัมพันธ์ของเนื้อหาที่ตรงกันหรือมีความคล้ายคลึงกับรูปแบบมิติความสัมพันธ์ของคำค้นของผู้ใช้ การนำเทคนิคการวิเคราะห์หมวดหมู่แบบสหความสัมพันธ์ (Multiple DDC Relational Classification) มาใช้จะทำให้สามารถวิเคราะห์หมวดหมู่สารสนเทศได้แม่นยำใกล้เคียงกับมนุษย์ซึ่งจะทำให้การดำเนินงานทำได้รวดเร็วขึ้นเสียค่าใช้จ่ายน้อยลง และการใช้กราฟิกแสดงผลหน้าในรูปแบบที่สอดคล้องจะทำให้เพิ่มประสิทธิภาพของเครื่องมือการค้นคืนและจะช่วยพัฒนาผู้ใช้ในการค้นคืนสารสนเทศ จากกรอบ

ความคิดที่นำเสนอผู้วิจัยจะดำเนินการศึกษาและพัฒนาต่อในเรื่องการจัดทำดัชนีเพื่อการค้นคืนแบบสหความสัมพันธ์และการวิเคราะห์หมวดหมู่แบบสหความสัมพันธ์ แล้วนำมาเปรียบเทียบความคล้ายคลึงด้วยฟัซซี่ ตลอดจนพัฒนาการแสดงผลด้วยกราฟิก เพื่อเพิ่มประสิทธิภาพในกระบวนการในการค้นคืนสารสนเทศต่อไป

7. เอกสารอ้างอิง

- [1] K. Markey. "Forty years of classification online: Final chapter or future unlimited," **Cataloging and Classification Quarterly**. Vol. 42, no. 3-4, pp.1-63, 2006.
- [2] Lois Mai Chan. **Cataloging and classification : An introduction**. 3rd.ed. Lanham, Md.: Scarecrow Press, 2007.
- [3] Mohinder Partap Satija. **The theory and practice of the Dewey Decimal Classification system**. Oxford: Chandos, 2007.
- [4] Mona L. Scott. "Dewey Decimal Classification", 22 nd edition: **A study manual and number building guide**. Westport, Conn.: Libraries Unlimited, 2005.
- [5] S. Dal Porto. and A. Marchitelli. "The functionality and flexibility of traditional classification schemes applied to a Content Management System (CMS): Facets, DDC, JITA". **Knowledge Organization**. vol. 34, no.1, pp. 35-44, 2006.
- [6] Wilaiporn Lertmahakiat and Anirach Mingkwan. "A propose idea of search engine results page base on DDC classification". **AITET 2nd conference: S&T Teaching in Vocational Education based on Sufficient Economy**. Vol.2, No.1, Jan-Dec, 2007.
- [7] Bill Kules, Jack Kustanowitz and Ben Shneiderman. "Categorizing web search results into meaningful and stable categories using fast-feature techniques," **ACM 1-59593-354-9/06/0006**, 2006.
- [8] Barbara Leporini, Andronico Patrizia and Buzzi, Marina. "Designing search engine user interfaces for the visually impaired," **ACM SIGCAPH Newsletter**.



- vol.18, no. 76(June), pp. 17-18, 2003.
- [9] R. Torres, C. Silva, C. Medeiros and H. Rocha. "Visual structures for image browsing," **CIKM'03**, New Orleans, Louisiana, November 3-8, 2003.
- [10] A. Spoerri. "How the overlap between the search results of different retrieval systems correlates with document relevance," **Proc. Of the 68th annual meeting of the American Society for Information Science and Technology (ASIST2005)**, 2005.
- [11] K. Golub. **Automated subject classification of textual web pages, based on a controlled vocabulary: challenges and recommendations**, Available on line at http://www.it.lth.se/koraljka/Lund/publ/Hypermedia_2006.pdf documents. retrieval 2007/09/21.
- [12] Keith F. Bourgoin. **Improving document classification through formal concept analysis and WordNet**. M.S. Thesis, Department of Computer Science Case Western Reserve University, 2006.
- [13] R. Prabowo, et al. "Ontology-based automatic classification for the Web". **Proceedings of the 3rd International conference on Web information systems engineering**, pp.182-191, 2002.
-
- ◆