

# การปรับปรุงการเข้าถึงผลลัพธ์การสืบค้นโดยการออกแบบ ส่วนของการแสดงผลการสืบค้นในเสิร์ชเอนจินด้วยเทคนิคการจัดกลุ่ม

## An Improvement of Search Results Access by Designing a Search Engine Result Page with a Clustering Technique

วรสิทธิ์ ชูชัยวัฒนา

ห้องปฏิบัติการวิจัยทางด้านเสิร์ชเอนจินและระบบสารสนเทศอัจฉริยะ  
วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอ็นเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิต กรุงเทพมหานคร  
worasit.cha@dpu.ac.th

### บทคัดย่อ

จากปัญหาข้อมูลเกินพิกัด ส่งผลทำให้เสิร์ชเอนจินกลายเป็นเครื่องมือที่สำคัญที่จะช่วยให้ผู้ใช้งานเว็บไซต์เว็บสามารถค้นหาข้อมูลตามความต้องการของตัวเองได้ งานวิจัยชิ้นนี้มีวัตถุประสงค์เสนอแนวคิดในการออกแบบหน้าจอการแสดงผลการสืบค้นด้วยเทคนิคการจัดกลุ่ม เพื่อที่จะส่งผลให้เพิ่มประสิทธิภาพในการเข้าถึงผลลัพธ์การสืบค้น โดยประยุกต์แนวคิดดังกล่าวกับการสืบค้นรูปภาพ ในการประเมินผลการออกแบบหน้าจอการแสดงผลการสืบค้น นักศึกษาระดับบัณฑิตศึกษาจากสาขาวิชาวิศวกรรมเว็บและการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา จำนวน 60 คนได้ถูกเชิญเข้ามาเป็นผู้ทดลองในงานวิจัย โดยนักศึกษาทุกคนจะถูกมอบหมายงานให้ทำการค้นรูปภาพ 2 ครั้ง จากระบบต้นแบบของเสิร์ชเอนจินที่มีการแสดงผลลัพธ์แบบลิสต์และแบบจัดกลุ่ม แบบแผนการทดลองจะใช้วิธี Within-Subjects Design ร่วมกับ Counter Balancing เพื่อลด Learning Effect โดยผลการทดลองสรุปได้ว่า การออกแบบหน้าจอแสดงผลด้วยเทคนิคการจัดกลุ่มสามารถเพิ่มประสิทธิภาพในการเข้าถึงผลลัพธ์การสืบค้น

**คำสำคัญ:** การออกแบบหน้าจอแสดงผลการสืบค้น, เสิร์ชเอนจิน, เทคนิคการจัดกลุ่ม

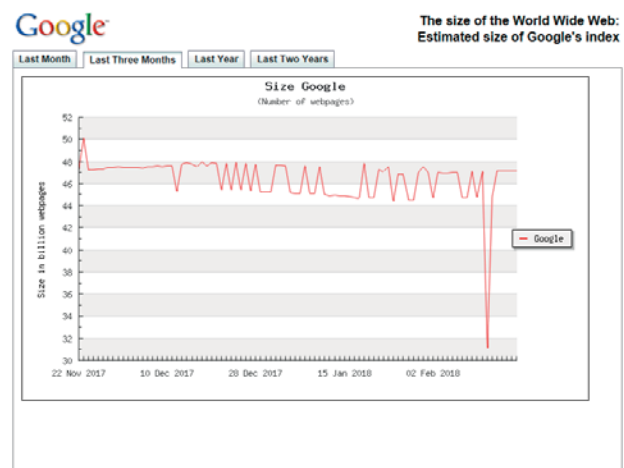
### ABSTRACT

Because of information overloaded problem, search engines become important tools to help WWW users to discover the information they desire. This paper aims at designing a search result page with a clustering technique to improve an efficiency of search results access by applying it to image searching task. To evaluate the proposed design, sixty graduate students from the department of Web Engineering and Mobile Application Development were recruited to be participants in this research study. They were given two tasks of finding images and were asked to use a prototype of search engine with displaying the results as lists and as clusters. The within-subjects design with a counter balancing technique for eliminating a learning effect was adopted as an experimental research design. The results showed that a search result page with a clustering technique could improve an efficiency of search results access.

**Keywords:** Search Result Page Design, Search Engine, Clustering Technique

### 1) บทนำ

ปริมาณข้อมูลบนอินเทอร์เน็ตและเว็บไซต์เว็บเพิ่มขึ้นจำนวนมหาศาล [1] เนื่องจากเป็นช่องทางที่ทำให้ผู้คนทั่วโลกสามารถสร้างและจัดเก็บข้อมูลต่าง ๆ นอกจากนั้นแล้วยังเป็นช่องทางในการแบ่งปันข้อมูลและการเข้าถึงข้อมูลต่าง ๆ ได้อย่างสะดวกและรวดเร็ว จากรายงาน [2] ของเว็บ [www.worldwidewebsite.com](http://www.worldwidewebsite.com) ในเดือนพฤศจิกายน 2560 ถึงเดือนมกราคม 2561 จำนวนเว็บเพจที่ถูกลีเกิลทำการสร้างดัชนีไว้มีมากกว่า 44,000 ล้านเว็บเพจ ดังรายละเอียดที่แสดงในรูปที่ 1



รูปที่ 1: ขนาดของเว็บไซต์เว็บประมาณการจากดัชนีของกูเกิล

ในการสืบค้นข้อมูลสารสนเทศบนอินเทอร์เน็ตและเว็บไซต์เว็บจึงกลายเป็นกิจกรรมประจำวันของคนทั่วโลก อย่างไรก็ตามการที่มีปริมาณข้อมูลมหาศาล ส่งผลทำให้เกิดปัญหาข้อมูลเกินพิกัด (Information Overload Problem) กับผู้ที่ต้องการสืบค้นข้อมูล [1] งานวิจัยส่วนใหญ่จึงให้ความสำคัญในเรื่องของการปรับปรุงกระบวนการทำงานของส่วนประกอบต่าง ๆ ของโปรแกรมการค้นหา หรือที่รู้จักกันในชื่อเสิร์ช เอนจิน (Search Engine) ไม่ว่าจะเป็นในเรื่องของการปรับปรุงกระบวนการรวบรวมเว็บเพจ (Webpage Crawling Process Improvement) การปรับปรุงกระบวนการสร้างดัชนีเว็บเพจ (Webpage

Indexing Process Improvement) และการปรับปรุงวิธีการเรียงลำดับผลลัพธ์การสืบค้น (Search Results Ranking Mechanism Improvement) โดยมีเป้าหมายเพื่อที่จะเพิ่มประสิทธิภาพและประสิทธิผลในการใช้งานเสิร์ชเอนจิน และสร้างประสบการณ์ที่ดีให้กับผู้ใช้งาน อย่างไรก็ตามงานวิจัยที่ให้ความสำคัญในเรื่องของการออกแบบส่วนแสดงผลการค้นหา (Search Engine Result Page) ยังมีอยู่ในปริมาณน้อย ดังนั้นงานวิจัยชิ้นนี้ จึงนำเสนอแนวทางในการออกแบบส่วนแสดงผลการค้นหาโดยใช้เทคนิคการจัดกลุ่ม (Clustering Technique)

## 2) ทบทวนวรรณกรรม

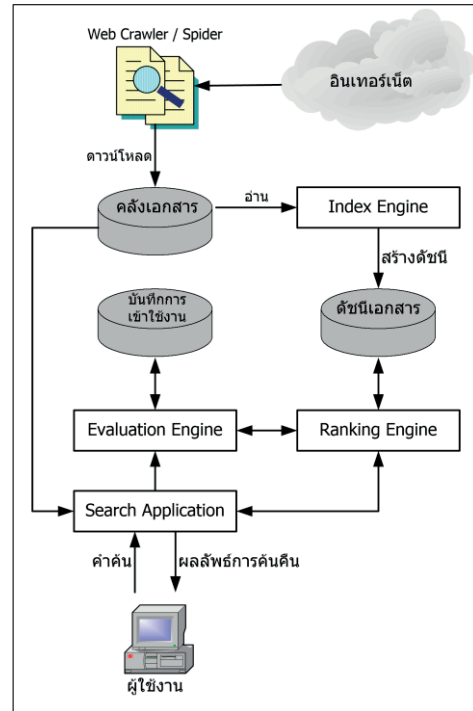
### 2.1) สถาปัตยกรรมของเสิร์ชเอนจิน

โดยทั่วไปแล้ว ระบบค้นคืนสารสนเทศและเสิร์ชเอนจิน จะประกอบไปด้วยส่วนประกอบที่สำคัญได้แก่ เว็บครอว์เลอร์ (Web Crawler) ส่วนสร้างดัชนี (Index Engine) แอปพลิเคชันสืบค้น (Search Application) ส่วนเรียงลำดับผลลัพธ์ (Ranking Engine) และ ส่วนการประเมินผลลัพธ์การสืบค้น (Evaluation Engine) [3] ดังแสดงรายละเอียดในรูปที่ 2 แต่ละส่วนประกอบจะมีหน้าที่และความรับผิดชอบในการทำงานที่แตกต่างกัน โดยมีจุดมุ่งหมายเดียวกัน เพื่อให้สามารถตอบสนองความต้องการให้กับผู้ใช้งานระบบได้อย่างมีประสิทธิภาพและมีประสิทธิผล

เว็บครอว์เลอร์ (Web Crawler) จะเป็นโปรแกรมขนาดเล็กที่มีหน้าที่ในการเข้าไปทำการดาวน์โหลดเว็บเพจมาเก็บไว้ที่คลังเอกสาร (Document Corpus) โดยเว็บเพจที่ถูกดาวน์โหลดมาจะถูกนำไปใช้งานในด้วยส่วนสร้างดัชนี (Index Engine) และแอปพลิเคชันสืบค้น (Search Application) ต่อไป

ส่วนสร้างดัชนี (Index Engine) รับผิดชอบในการนำเว็บเพจในคลังเอกสารไปทำการสกัดเนื้อหา เพื่อให้ได้เนื้อหาที่เป็นตัวแทนของเอกสาร และทำการสร้างดัชนีของเอกสารซึ่งจะส่งผลให้เพิ่มความเร็วในการทำงานของแอปพลิเคชันสืบค้น (Search Application) ในส่วนนี้จะเป็นส่วนที่จะโต้ตอบกับผู้ใช้งานระบบ โดยมีหน้าที่รับคำค้น (Query) ปรับปรุงคำค้น (Query Refinement) และแสดงผลการค้นหา (Search Results) เมื่อแอปพลิเคชันสืบค้นได้รับคำค้น จะทำการส่งคำค้นไปยังส่วนเรียงลำดับผลลัพธ์ (Ranking Engine) ซึ่งจะทำหน้าที่ในการนำคำค้นไปเปรียบเทียบกับดัชนีของเอกสาร เพื่อคำนวณความเกี่ยวพัน (Relevant) ระหว่างคำค้นและดัชนีเอกสาร จากนั้นจะทำหน้าที่ในการเรียงลำดับผลลัพธ์ของการสืบค้น ก่อนที่จะส่งผลลัพธ์ดังกล่าวกลับมายังแอปพลิเคชันสืบค้น เพื่อแสดงผลการค้นหาให้กับผู้ใช้

นอกจากนั้นแล้ว ส่วนการประเมินผลลัพธ์การสืบค้น (Evaluation Engine) จะทำงานร่วมกับแอปพลิเคชันสืบค้น (Search Application) ในการบันทึกผลลัพธ์การสืบค้น และพฤติกรรมในการพิจารณาผลลัพธ์ของผู้ใช้ เพื่อที่จะนำข้อมูลดังกล่าวไปใช้ในการปรับปรุงประสิทธิภาพและประสิทธิผลของเสิร์ชเอนจินต่อไป



รูปที่ 2: สถาปัตยกรรมของระบบค้นคืนสารสนเทศและเสิร์ชเอนจิน

### 2.2) งานวิจัยที่เกี่ยวข้อง

ในการปรับปรุงประสิทธิภาพและประสิทธิผลของเสิร์ชเอนจิน [1] สามารถทำได้หลายแนวทางขึ้นอยู่กับมุมมองของนักวิจัยว่าต้องการจะปรับปรุงการทำงานในส่วนประกอบใดของเสิร์ชเอนจิน โดยอาจจะเป็นการปรับปรุงกระบวนการในการสร้างดัชนีตัวแทนเอกสาร และการปรับปรุงเทคนิคที่ใช้ในการเรียงลำดับผลลัพธ์การสืบค้น

การสร้างดัชนีที่เป็นตัวแทนของเอกสารได้ดี จะส่งผลให้ผลลัพธ์การสืบค้นตรงกับความต้องการของผู้สืบค้นมากขึ้น ดังนั้นนักวิจัยทางด้านเสิร์ชเอนจินและทางด้านสารสนเทศศาสตร์ได้เสนอแนวคิดและพัฒนาวิธีการต่าง ๆ เพื่อปรับปรุงคุณภาพของดัชนีของเอกสาร โดยอาจจะเป็นการปรับปรุงคุณภาพของดัชนีด้วยวิธีการวิเคราะห์โครงสร้างเอกสาร ได้แก่ การสกัดโครงสร้างของเอกสารในส่วนของการเชื่อมโยงเอกสาร (Document Title) และส่วนของข้อความลิงค์เชื่อมโยง (Anchor Text) [4-7] นอกจากนั้นแล้ว มีนักวิจัยอีกกลุ่มใช้วิธีการวิเคราะห์บันทึกการใช้งานของผู้ใช้ (User Usage Log) มาใช้ในการปรับปรุงคุณภาพของดัชนี โดยงานวิจัยในแนวนี้ ซึ่งเป็นการใช้ประโยชน์จากส่วนการประเมินผลลัพธ์การสืบค้น (Evaluation Engine) ซึ่งทำหน้าที่เก็บคำค้น (Query) ผลลัพธ์การสืบค้น (Search Results List) และพฤติกรรมการเข้าสู่ผลลัพธ์ของผู้ใช้ แม้ว่าพฤติกรรมการเข้าสู่ผลลัพธ์ของผู้ใช้จะไม่สามารถรับประกันได้ว่าผลลัพธ์ดังกล่าวจะมีความเกี่ยวพัน (Relevant) กับคำค้น แต่การพิจารณาข้อมูลดังกล่าวก็สามารถทำให้ประเมินได้ว่าผลลัพธ์ใดจะเป็นผลลัพธ์ที่น่าสนใจของคำค้นนั้น ๆ โดยสามารถนำข้อมูลดังกล่าวไปใช้สำหรับการปรับปรุงดัชนีของเอกสารและวิธีการการเรียงลำดับผลลัพธ์การสืบค้นได้ [8-9] ในขณะเดียวกันมีนักวิจัยอีกจำนวนหนึ่งให้ความสนใจในการนำเอาข้อมูลอื่น ๆ นอกเหนือจากที่ปรากฏในเอกสารมาใช้ประโยชน์ เช่น ข้อมูลจากบรรณนิทัศน์เชิงสังคม (Social Annotation) หรือ ป้ายระบุข้อมูลเชิงสังคม

(Social Tag) มาใช้เพื่อปรับปรุงคุณภาพของการสืบค้นข้อมูล โดยการนำไปใช้ในการสร้างดัชนีที่เป็นตัวแทนของเอกสาร [10-13]

การพิจารณาคุณภาพของเอกสารเพื่อปรับปรุงวิธีการเรียงลำดับผลลัพธ์การสืบค้น เป็นอีกแนวทางหนึ่งที่นักวิจัยให้ความสนใจ Brin and Page ผู้ก่อตั้งกูเกิลได้เสนอแนวคิดวิธีการเรียงลำดับผลลัพธ์การสืบค้น โดยจะพิจารณาคุณภาพของเอกสาร [4] ซึ่งแนวทางดังกล่าวได้กลายมาเป็นส่วนหนึ่งของวิธีการเรียงลำดับผลลัพธ์การสืบค้นของกูเกิล ซึ่งเป็นเสิร์ชเอนจินที่มีคนทั่วโลกใช้มากที่สุดตัวหนึ่ง และงานวิจัยต่อ ๆ มา ก็ได้นำเอาแนวคิดดังกล่าวพัฒนาเพิ่มเติม ซึ่งปรากฏอยู่ในงานของ Choochaiwattana and Spring [10] และงานของ Jomsri et al. [14]

ใน [15] มีการสำรวจและวิเคราะห์รูปแบบการแสดงผลการสืบค้นของเสิร์ชเอนจินที่เป็นที่นิยมจำนวนทั้งสิ้น 12 ตัว และได้ทำการสรุปองค์ประกอบที่สำคัญของรูปแบบการแสดงผลการสืบค้น โดยได้ระบุไว้ว่าการแบ่งกลุ่มเอกสารที่เป็นผลลัพธ์ของการสืบค้นนั้น จะทำให้ผู้ค้นหาสามารถตรวจสอบผลลัพธ์การสืบค้นได้เร็วขึ้น ดังนั้นในงานวิจัยชิ้นนี้จึงนำเอาประเด็นดังกล่าวมาพัฒนาเพิ่มเติม โดยจะให้ความสำคัญในเรื่องของการออกแบบที่จะทำให้สามารถเพิ่มประสิทธิภาพของการเข้าถึงผลลัพธ์ของการสืบค้น ซึ่งจะเป็นแนวทางที่สำคัญที่จะส่งผลทำให้เกิดการปรับปรุงประสิทธิภาพและประสิทธิผลของเสิร์ชเอนจินต่อไปในอนาคต

### 3) วัตถุประสงค์และสมมติฐานของงานวิจัย

วัตถุประสงค์ของงานวิจัยนี้ เพื่อทำการออกแบบรูปแบบการแสดงผลการสืบค้นโดยการประยุกต์ใช้เทคนิคการจัดกลุ่ม (Clustering Technique) โดยจะนำเอาแนวทางในการออกแบบดังกล่าวไปประยุกต์ใช้กับการพัฒนาเสิร์ชเอนจินต้นแบบสำหรับการสืบค้นรูปภาพบนเว็บไซต์ไวด์เว็บ (Prototype of WWW Image Search Engine) และทำการประเมินประสิทธิภาพของรูปแบบการแสดงผลการสืบค้นดังกล่าว โดยเปรียบเทียบกับประสิทธิภาพของการแสดงผลรูปแบบปกติของการสืบค้นรูปภาพบนเสิร์ชเอนจิน

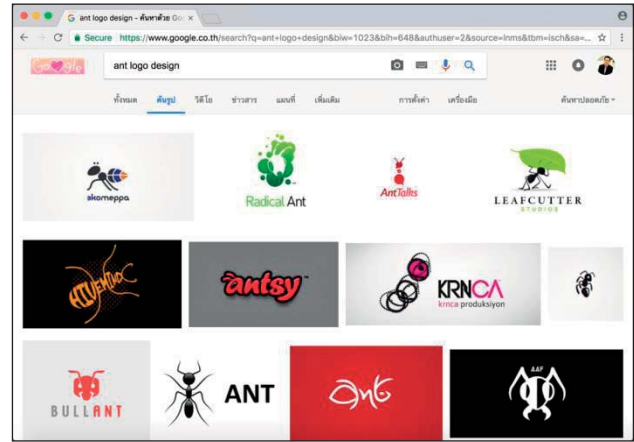
โดยกำหนดสมมติฐานของงานวิจัยไว้ว่า รูปแบบการแสดงผลการสืบค้นรูปภาพแบบที่มีการใช้เทคนิคการจัดกลุ่มจะให้ประสิทธิภาพในการเข้าถึงผลลัพธ์การสืบค้นมากกว่ารูปแบบการแสดงผลการสืบค้นรูปแบบแบบปกติที่เสิร์ชเอนจินมีการใช้อยู่ในปัจจุบัน โดยพิจารณาจากเวลาที่ใช้ในการเข้าถึงผลลัพธ์ของการสืบค้น

### 4) ระเบียบวิธีวิจัยและการออกแบบการทดลอง

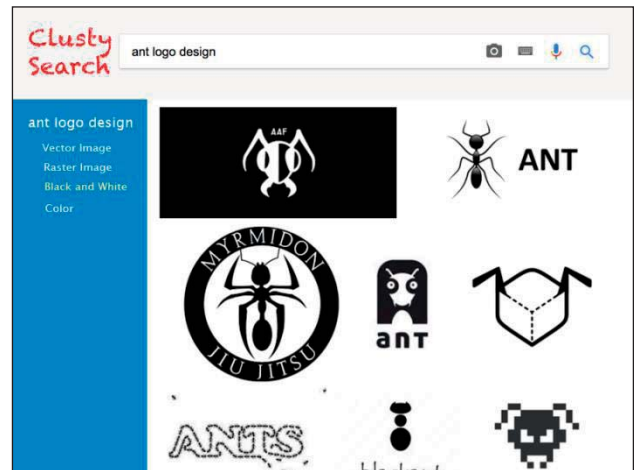
#### 4.1) การออกแบบการแสดงผลการสืบค้นด้วยเทคนิคการจัดกลุ่ม

การออกแบบการแสดงผลการสืบค้นที่ง่ายต่อการใช้งานจะส่งผลทำให้สามารถสร้างประสบการณ์ที่ดีให้กับผู้ใช้เสิร์ชเอนจินได้ โดยปกติแล้วการแสดงผลการสืบค้นบนเสิร์ชเอนจินจะแสดงในรูปแบบของลิสต์ของผลลัพธ์การสืบค้น (Search Results List) ดังแสดงในรูปที่ 3 ซึ่งจะเห็นได้ว่าผลการแสดงผลการสืบค้นแบบดังกล่าว ผู้ใช้งานเสิร์ชเอนจินมักจะใช้เวลาในการพิจารณาผลลัพธ์การสืบค้นค่อนข้างมาก ดังนั้นในงานวิจัยนี้ จึงนำเอาเทคนิคการจัดกลุ่ม (Clustering Technique) มาใช้เพื่อการจัดกลุ่มผลลัพธ์การสืบค้น เพื่อเพิ่มช่องทาง

ให้กับผู้ใช้ในการเลือกกลุ่มของผลลัพธ์ตามหัวข้อที่เกี่ยวข้องกับผลลัพธ์ที่แสดงในรูปที่ 4 โดยในด้านซ้ายมือของหน้าจอ จะทำการแสดงการจัดกลุ่มของลักษณะของผลลัพธ์การสืบค้น



รูปที่ 3: การออกแบบการแสดงผลการสืบค้นรูปภาพบนกูเกิล



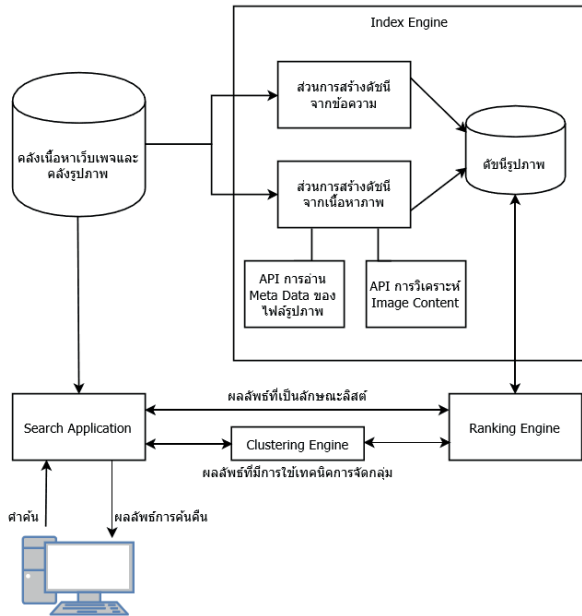
รูปที่ 4: การออกแบบการแสดงผลการสืบค้นรูปภาพที่มีการใช้เทคนิคการจัดกลุ่ม

#### 4.2) การเตรียมข้อมูลและการพัฒนาเสิร์ชเอนจินต้นแบบ

เว็บเบราว์เซอร์ทำการดาวน์โหลดเว็บเพจและรูปภาพมาเก็บไว้ในคลังเอกสารในช่วงระหว่างเดือนมกราคมถึงเดือนกันยายน 2560 อย่างไม่จำกัด เนื่องจากวัตถุประสงค์ของงานวิจัยนี้เน้นที่รูปแบบของการแสดงผลการสืบค้น ดังนั้นเว็บเบราว์เซอร์ที่ได้ถูกพัฒนาขึ้น จะทำการดาวน์โหลดเนื้อหาเว็บเพจและรูปภาพ โดยจะนำเอาเนื้อหาของเว็บเพจ ร่วมกับการใช้ API ในการอ่าน Metadata ของไฟล์รูปภาพ และการวิเคราะห์ Image Content ไปสร้างเป็นดัชนีตัวแทนรูปภาพ เพื่อใช้สำหรับแอปพลิเคชันสืบค้น และในการสร้างดัชนีนั้นจะใช้ตัวแบบเวกเตอร์สเปซโมเดล [3] ซึ่งเป็นตัวแบบมาตรฐานสำหรับการสร้างดัชนีในระบบค้นคืนสารสนเทศ โดยใช้ TF-IDF เป็นการกำหนดค่าน้ำหนักของแต่ละเทอมในเอกสาร และการเรียงลำดับผลลัพธ์การสืบค้นของส่วนเรียงลำดับผลลัพธ์ (Ranking Engine) จะใช้ Cosine Similarity เปรียบเทียบคำค้น (Query) และดัชนีของดัชนีรูปภาพ ดังที่แสดงในสมการที่ 1

$$Sim(D_j, Q) = \frac{\sum_{i=1}^N w_{i,j} \cdot w_{i,q}}{\sqrt{\sum_{i=1}^N w_{i,j}^2 \cdot \sum_{i=1}^N w_{i,q}^2}} \quad (1)$$

เพื่อวัดประสิทธิภาพของการเข้าถึงผลลัพธ์การสืบค้น จากการออกแบบหน้าจอแสดงผลการสืบค้นด้วยเทคนิคการจัดกลุ่มเสิร์ชเอนจินต้นแบบ (Search Engine Prototype) จึงได้ถูกพัฒนาขึ้น โดยประกอบไปด้วยส่วนของการแสดงผลการสืบค้น 2 รูปแบบ ได้แก่ การแสดงผลลัพธ์ที่เป็นลักษณะลิสต์ และการแสดงผลลัพธ์ที่มีการใช้เทคนิคการจัดกลุ่ม ดังรูปที่ 5 ซึ่งแสดงส่วนประกอบของเสิร์ชเอนจินต้นแบบ



รูปที่ 5 : ส่วนประกอบของเสิร์ชเอนจินต้นแบบ

สำหรับในส่วนของ Clustering Engine นั้น ผู้วิจัยเลือกใช้เทคนิค Hierarchical Clustering โดยใช้ค่านำหนักของเทอมของดัชนีรูปภาพเป็นตัวเปรียบเทียบเพื่อจัดกลุ่มของรูปภาพ ในขณะที่เดียวกันการตั้งชื่อกลุ่ม (Cluster Labeling) จะทำโดยการพิจารณาจากความถี่ของเทอมของดัชนีรูปภาพในกลุ่มเดียวกัน

#### 4.3) การออกแบบการทดลองและการประเมินผลการออกแบบ

ในการประเมินผลการออกแบบ นักศึกษาระดับบัณฑิตศึกษาจากสาขาวิศวกรรมเว็บและการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา จำนวน 60 คน เข้ามาเป็นผู้ทดลอง โดยแต่ละท่านได้รับคำสั่งให้ทำการค้นหาข้อมูลที่เกี่ยวข้องกับรูปสินค้าที่ได้กำหนดไว้ล่วงหน้า 2 รูป โดยผู้ทดลองจะเห็นรูปภาพดังกล่าว และให้ทำการกำหนดคำค้น (Query) ด้วยตนเอง เมื่อเสิร์ชเอนจินต้นแบบรับคำค้นดังกล่าวแล้ว จะทำการสืบค้นและแสดงผลการสืบค้นให้กับผู้ทดลอง โดยผู้ควบคุมการทดลองจะทำการบันทึกเวลา ตั้งแต่เริ่มส่งคำค้นเข้าสู่เสิร์ชเอนจินต้นแบบจนกระทั่งผู้ทดลองค้นพบรูปภาพตามคำสั่งที่ระบุไว้

เนื่องจากลักษณะของส่วนแสดงผลการสืบค้นมี 2 รูปแบบ แต่มีผู้ทำการทดลองแค่ 1 กลุ่ม ดังนั้นเพื่อลดผลกระทบจากการเรียนรู้ (Learning Effects) ของการใช้งานเสิร์ชเอนจินต้นแบบ แบบแผนการ

ทดลองจึงจำเป็นต้องใช้วิธี Within-Subjects Design ร่วมกับ Counter Balancing โดยการแสดงผลการสืบค้นจะสลับกันโดยที่ผู้ทดลอง 30 คนแรก ในการทำการทดลองครั้งที่ 1 ส่วนแสดงผลการสืบค้นจะเป็นแบบลิสต์ก่อน แล้วจะค่อยแสดงผลลัพธ์ที่เป็นแบบที่มีการใช้เทคนิคการจัดกลุ่ม และจะสลับกันในการทำการทดลองครั้งที่ 2 ในขณะที่ผู้ทดลอง 30 คนหลัง ในการทำการทดลองครั้งที่ 1 ส่วนแสดงผลการสืบค้นแบบที่มีการใช้เทคนิคการจัดกลุ่มจะแสดงก่อน แล้วค่อยแสดงผลลัพธ์ที่เป็นแบบลิสต์ และจะสลับกันในการทำการทดลองครั้งที่ 2

จากนั้นจะนำเอาเวลาที่บันทึกไว้ของผู้ทดลองทุกคนมาหาค่าเฉลี่ย เพื่อเปรียบเทียบเวลาที่ใช้ในการเข้าถึงผลลัพธ์การสืบค้นที่ถูกต้องของส่วนแสดงผลการสืบค้นทั้งสองแบบ นอกจากนั้นแล้ว ผู้ควบคุมการทดลองจะทำการสัมภาษณ์ผู้ทดลองอย่างไม่เป็นทางการในประเด็นที่เกี่ยวข้องกับผลกระทบของการออกแบบส่วนแสดงผลการสืบค้น เพื่อสรุปความคิดเห็นและความรู้สึกของการใช้งานส่วนแสดงผลการสืบค้นแบบที่มีการใช้เทคนิคการจัดกลุ่ม

#### 5) ผลการวิจัย การอภิปรายผล และข้อเสนอแนะ

เนื่องจากผู้ทดลองแต่ละคนจะได้เห็นรูปภาพที่เป็นผลลัพธ์ก่อนที่จะเริ่มทำการสืบค้นผ่านเสิร์ชเอนจินต้นแบบ และเปิดโอกาสให้ผู้ทดลองทำการกำหนดคำค้นด้วยตนเอง ดังนั้นในการจับเวลาในการเข้าถึงผลลัพธ์การสืบค้นจะเริ่มตั้งแต่ผู้ทดลองส่งคำค้นเข้าไปในเสิร์ชเอนจินต้นแบบ จนกระทั่งผู้ทดลองสามารถเข้าถึงผลลัพธ์ที่ถูกต้องตรงกับรูปภาพในคำสั่ง ตารางที่ 1 แสดงเวลาเฉลี่ยที่ผู้ทดลองใช้สำหรับการเข้าถึงผลลัพธ์การสืบค้น ผ่านส่วนแสดงผลการสืบค้นทั้งแบบลิสต์และแบบที่มีการใช้เทคนิคการจัดกลุ่ม ซึ่งจะพบว่าเวลาเฉลี่ยของแบบที่มีการใช้เทคนิคการจัดกลุ่มน้อยกว่าแบบลิสต์ 11.15 วินาที ซึ่งสามารถสรุปได้ว่าผู้ทดลองเข้าถึงผลลัพธ์การสืบค้นได้เร็วกว่าเมื่อใช้งานส่วนการแสดงผลการสืบค้นที่มีการใช้เทคนิคการจัดกลุ่ม อย่างไรก็ตามส่วนเบี่ยงเบนมาตรฐานของส่วนการแสดงผลการสืบค้นที่มีการใช้เทคนิคการจัดกลุ่มมากกว่า

ตารางที่ 1: แสดงเวลาเฉลี่ยของการเข้าถึงผลลัพธ์การสืบค้นของผู้ทดลอง

รูปแบบการแสดงผลการสืบค้น	ผลการประเมิน	
	เวลาเฉลี่ย (วินาที)	ส่วนเบี่ยงเบนมาตรฐาน
แบบลิสต์ (List)	56.24	9.28
แบบที่มีการใช้เทคนิคการจัดกลุ่ม (Clustering Technique)	45.09	12.44

หากพิจารณาถึงสาเหตุหลักที่ทำให้การเข้าถึงผลลัพธ์การสืบค้นของรูปแบบการแสดงผลการสืบค้นแบบจัดกลุ่มมีความรวดเร็วกว่ารูปแบบการแสดงผลแบบลิสต์ เป็นเพราะว่าพฤติกรรมของผู้สืบค้นมักจะใช้คำค้น (Query Term) ที่มีความหมายค่อนข้างกว้าง ในกรณีนี้ที่เพิ่ม

เริ่มทำการสืบค้น การใช้คำค้นที่ไม่เฉพาะเจาะจง จะส่งผลโดยตรงต่อ เวลาที่ใช้ในการพิจารณาผลลัพธ์การสืบค้น ดังนั้นการจัดรูปแบบการ แสดงผลลัพธ์การสืบค้นแบบกลุ่ม จะช่วยให้ผู้สืบค้นสามารถเข้าถึง กลุ่มผลลัพธ์ที่ตัวเองต้องการได้ทันที จึงส่งผลให้ลดระยะเวลาในการ เข้าถึงผลลัพธ์ที่ต้องการได้

เมื่อวิเคราะห์รายละเอียดของข้อมูลเวลาในการเข้าถึงผลลัพธ์การ สืบค้นของส่วนแสดงผลทั้งสองแบบพบว่า มีผู้ทดลองอยู่ส่วนหนึ่งมี ความไม่คุ้นเคยกับการแสดงผลที่มีการจัดกลุ่ม (Clustering) จะใช้ เวลาในการเข้าถึงผลลัพธ์มากกว่าแบบลิสต์ แต่เมื่อเข้าใจการทำงานของ ส่วนแสดงผลการสืบค้นแบบจัดกลุ่มแล้ว ก็สามารถใช้งานได้ รวดเร็วมากขึ้น ซึ่งสอดคล้องกับข้อมูลที่ได้จากการสัมภาษณ์อย่างไม่ เป็นทางการหลังการทำการทดลอง โดยผู้ทดลองจำนวนหนึ่งให้ ความเห็นว่า การเข้าถึงผลลัพธ์การสืบค้นแบบจัดกลุ่มสามารถทำให้ เข้าถึงผลลัพธ์ที่ต้องการได้ทันที แต่จะต้องทำความเข้าใจในส่วนของ การแสดงผลก่อน และเมื่อให้เปรียบเทียบส่วนแสดงผลการ สืบค้นทั้งสองแบบ ผู้ทดลองร้อยละ 75 มีความพึงพอใจกับการแสดงผล การสืบค้นแบบจัดกลุ่ม โดยให้เหตุผลว่าการจัดกลุ่มทำให้สามารถ เข้าไปถึงผลลัพธ์ได้ทันที ทำให้สะดวกในการเข้าถึงผลลัพธ์ได้มากกว่า อย่างไม่รู้ดี มีผู้ทดลองบางท่านให้ข้อเสนอแนะว่า ข้อความที่ใช้เป็น คำอธิบายกลุ่ม จะต้องมีความชัดเจนด้วย ซึ่งจะส่งผลทำให้การจัดกลุ่มมี ประสิทธิภาพและมีประสิทธิภาพมากยิ่งขึ้น

แม้ว่างานวิจัยชิ้นนี้ทำการทดลองบนเสิร์ชเอนจินต้นแบบที่สืบค้น รูปภาพ ผู้พัฒนาเสิร์ชเอนจินที่ต้องการเพิ่มประสิทธิภาพในการเข้าถึง ผลลัพธ์การสืบค้นสามารถนำแนวทางดังกล่าวไปใช้ในพัฒนา เสิร์ชเอนจินบนเอกสารชนิดอื่น ๆ เช่น เว็บเพจ เป็นต้น ได้เช่นเดียวกัน อย่างไม่รู้ดี วิธีการจัดกลุ่มเพื่อแสดงผลการสืบค้นอาจจะต้องมี การศึกษาเพิ่มเติม เพื่อความเหมาะสมกับชนิดของเอกสาร

ในอนาคตการศึกษาเพิ่มเติมในเรื่องเทคนิคการจัดกลุ่มเพื่อใช้ในการ แสดงผลลัพธ์การสืบค้น และการสร้างคำอธิบายกลุ่มจำเป็นต้อง ดำเนินการต่อไป รวมทั้งทดลองการออกแบบส่วนของการแสดงผล การสืบค้นในรูปแบบอื่น ๆ ที่จะเพิ่มประสิทธิภาพและประสิทธิผล ให้กับผู้ใช้เสิร์ชเอนจินในการเข้าถึงข้อมูลต่าง ๆ จำเป็นต้องศึกษา เพิ่มเติม

## 6) สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อทำการออกแบบรูปแบบการแสดงผล การสืบค้นโดยการประยุกต์ใช้เทคนิคการจัดกลุ่ม (Clustering Technique) มาใช้เพื่อปรับปรุงประสิทธิภาพของการเข้าถึงผลลัพธ์การ สืบค้น เพื่อประเมินประสิทธิภาพของการออกแบบ เสิร์ชเอนจินต้นแบบ จึงถูกพัฒนาขึ้นโดยประกอบไปด้วยส่วนของการแสดงผล 2 รูปแบบ ได้แก่ การแสดงผลที่เป็นลักษณะลิสต์ และการแสดงผลที่มีการใช้เทคนิคการจัดกลุ่ม ในการประเมินผลการออกแบบ นักศึกษาระดับบัณฑิตศึกษาจากสาขาวิชาวิศวกรรมเว็บและการพัฒนา แอปพลิเคชันบนอุปกรณ์พกพาจำนวน 60 คน เข้ามาเป็นผู้ทดลอง ผล การทดลองพบว่าเวลาเฉลี่ยในการเข้าถึงผลลัพธ์การสืบค้นของแบบที่ มีการใช้เทคนิคการจัดกลุ่มน้อยกว่าแบบลิสต์ 11.15 วินาที ซึ่งสามารถ สรุปได้ว่าผู้ทดลองเข้าถึงผลลัพธ์การสืบค้นได้รวดเร็วกว่าเมื่อใช้งาน

ส่วนการการแสดงผลการสืบค้นที่มีการใช้เทคนิคการจัดกลุ่ม และ จากการสัมภาษณ์ผู้ทดลองหลังจากการทำงานประเมินผลการออกแบบ พบว่า ผู้ทดลองร้อยละ 75 มีความพึงพอใจกับการแสดงผลแบบจัด กลุ่ม โดยให้เหตุผลว่าการจัดกลุ่มทำให้สามารถเข้าถึงผลลัพธ์ได้ โดยตรง แต่ได้ให้ข้อเสนอแนะว่า ข้อความที่ใช้ในการอธิบายกลุ่มจะต้อง มีความชัดเจนด้วย จึงจะส่งผลทำให้ประสิทธิภาพของการเข้าถึงผลลัพธ์ การสืบค้นดีขึ้น

## เอกสารอ้างอิง

- [1] วรสิทธิ์ ชูชัยวัฒนา, "การปรับปรุงประสิทธิภาพของระบบค้นคืนสารสนเทศและ โปรแกรมการค้นหา : แนวคิดและเทคนิค," *วารสารวิชาการสถาบันอุดมศึกษา เอกชนแห่งประเทศไทย ฉบับวิทยาศาสตร์และเทคโนโลยี*, ปีที่ 3, ฉบับที่ 1, หน้า 73 – 83, 2557
- [2] "WorldWideWebSize.com | The size of the World Wide Web (The Internet)." [Online]. Available: <http://www.worldwidewebsite.com/>. [Accessed: 14-Feb-2018].
- [3] B. Croft, D. Metzler, and T. Strohan, *Search Engines: Information Retrieval in Practice*. Boston: Pearson, 2009.
- [4] S. Brin and L. Page, "The Anatomy of a Large-scale Hypertextual Web Search Engine," in *Proceedings of the Seventh International Conference on World Wide Web 7*, Amsterdam, The Netherlands, The Netherlands, 1998, pp. 107–117.
- [5] N. Craswell, D. Hawking, and S. Robertson, "Effective Site Finding Using Link Anchor Information," in *Proceedings of the 24<sup>th</sup> Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, NY, USA, 2001, pp. 250–257.
- [6] T. H. W. Westerveld, W. Kraaij, and D. Hiemstra, "Retrieving Web Pages using Content, Links, URLs and Anchors," in *Proceedings of the Tenth Text retrieval Conference (TREC 2001)*, Gaithersburg, USA, 2001, pp. 663–672
- [7] A. Kundu, S. Sett, S. Kumar, S. Sengupta, and S. Chaudhury, "Search Engine Indexing Storage Optimization Using Hamming Distance," *International Journal of Intelligent Information and Database Systems*, vol. 6, no. 2, pp. 113–128, Mar. 2012.
- [8] G. Xue, H. Zeng, Z. Chen, T. Yu, W. Ma, W. Xi, and W. Fan, "Optimizing Web Search Using Web Click-Through Data," *Proceedings of the 13<sup>th</sup> ACM Conference on Information and Knowledge Management*, Washington DC, USA, 2004, pp.118-126.
- [9] F. Ahmadi-Abkenari and A. Selamat, "Advantages of Employing Log Rank Web Page Importance Metric in Domain Specific Web Search Engines," *International Journal of Digital Content Technology and its Application*, vol. 7, no. 9, pp. 425-432, 2013.
- [10] W. Choochaiwattana and M. Spring, "Applying Social Annotations to Retrieve and Re-rank Web Resources," *Proceedings of the International Conference on Information Management and Engineering*, Kuala Lumpur, Malaysia, 2009, pp. 215-219.
- [11] P. Dmitriev, N. Eiron, M. Fontoura, and E. Shekita, "Using Annotations in Enterprise Search," *Proceedings of the 15<sup>th</sup> International World Wide Web Conference*, Edinburgh, Scotland, 2006, pp. 811-817.
- [12] C. Marlow, M. Naaman, D. Boyd, and A. Davis, "Tagging Paper, Taxonomy, Flickr, Academic Article, To Read," *Proceedings of the 17<sup>th</sup> ACM Conference on Hypertext and Hypermedia*, Odense, Denmark, 2006, pp. 31-40.



- [13] D. Zhou, J. Bian, S. Zheng, H. Zha, and C.L. Giles, "Exploring Social Annotations for Information Retrieval," *Proceedings of the 17<sup>th</sup> International World Wide Web Conference*, Beijing, China, 2008, pp. 715-724.
- [14] P. Jomsri, S. Sanguansintukul, and W. Choochaiwattana, "CiteRank: Combination Similarity and Static Ranking with Research Paper Searching," *International Journal of Internet Technology and Secured Transactions*, vol. 3, no. 2, pp. 161-177, 2011.
- [15] วิไลพร เลิศมหาเกียรติ, ภูริวัตร คัมภีร์ภาพพัฒน และ อนิราช มิ่งขวัญ, "รูปแบบการแสดงผลการค้นคืนของเครื่องมือการสืบค้นสารนิเทศบนอินเทอร์เน็ต," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 18, ฉบับที่ 1, หน้า 89 – 98, 2551