

# ตัวแบบสำหรับการคัดเลือกข้อความบนหน้าเว็บเพจเพื่อการสร้างดัชนีของรูปภาพบนเว็บ โดยใช้ตัวแบบของเอกสาร

## A Model of Content Selection on Web Pages for Web Image Indexing Using Document Object Modeling

วิญญู นิรนาทล้ำพงศ์

วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอ็นเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิต

110/1-4 ถนนประชาชื่น เขตหลักสี่ กรุงเทพฯ 10210

E-mail: winyu.nir@dpu.ac.th

### บทคัดย่อ

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อที่จะนำเสนอตัวแบบสำหรับการคัดเลือกข้อความบนหน้าเว็บเพจเพื่อการสร้างดัชนีรูปภาพบนเว็บโดยใช้ตัวแบบเอกสาร ตัวแบบที่เสนอมจะมีแนวทางในการคัดเลือกทั้งหมด 4 แนวทางได้แก่ ALT, ImgSlibling, AllSubtree และ ParentSibling ทั้ง 4 แนวทางดังกล่าวเป็นการใช้หลักการวิเคราะห์ตัวแบบเอกสารโดยพิจารณาข้อความที่ใกล้กับตำแหน่งรูปภาพมากที่สุด และค่อยๆ ขยับตำแหน่งออกไปในส่วนที่ไกลออกไป ระบบค้นคืนรูปภาพต้นแบบได้ถูกพัฒนาขึ้นเพื่อใช้สำหรับการประเมินประสิทธิภาพของตัวแบบ โดยกำหนดสมมติฐานไว้ว่า ถ้าตัวแบบคัดเลือกข้อความได้เหมาะสม จะทำให้ได้ดัชนีรูปภาพที่ดี และดัชนีรูปภาพที่ดี จะส่งผลทำให้ผลลัพธ์การค้นคืนรูปภาพที่ดีด้วย จากผลการทดลองพบว่า ผลลัพธ์การค้นคืนที่ได้จากแนวทาง AllSubtree มีค่าความแม่นยำเฉลี่ยสูงที่สุด ได้แก่ 0.74 ดังนั้นสามารถสรุปได้ว่าแนวทางที่เสนอในงานวิจัยชิ้นนี้ มีความเหมาะสม และสามารถ

นำไปพัฒนาให้เป็นกระบวนการทำงานส่วนหนึ่งของเครื่องสร้างดัชนีรูปภาพของระบบค้นคืนรูปภาพได้

### Abstract

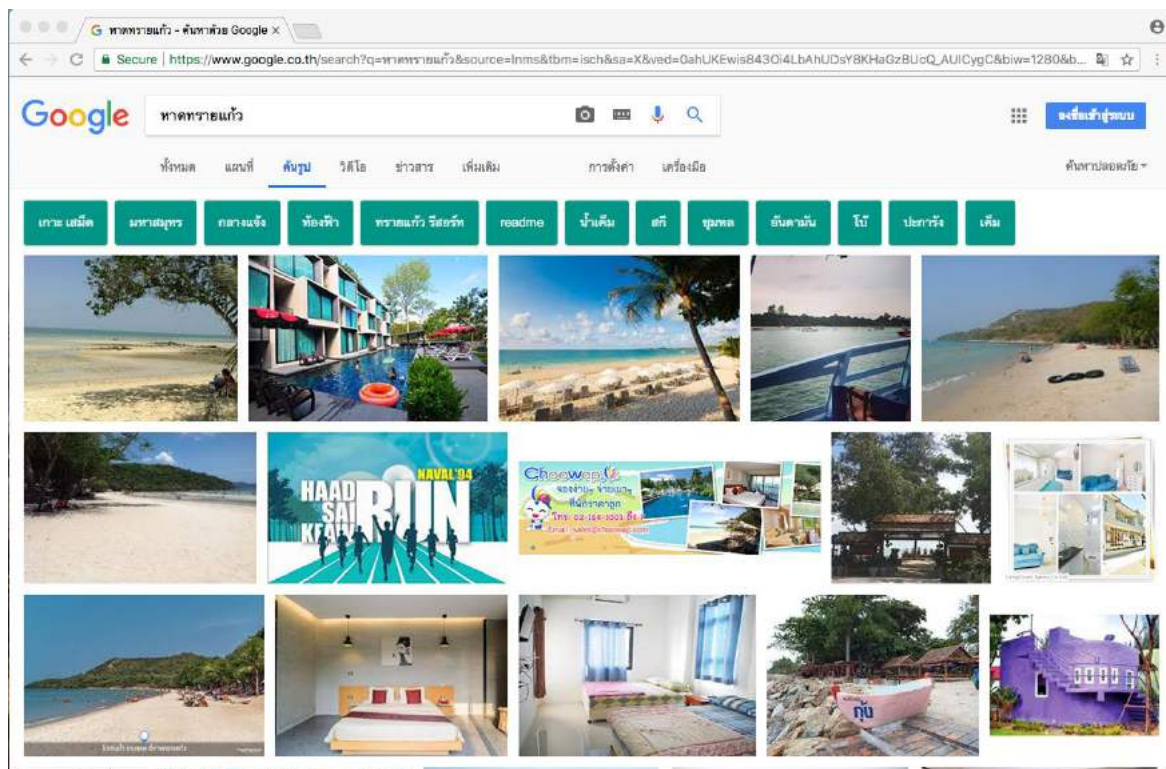
This research aims at proposing a model of content selection on web pages for web image indexing using Document Object Modeling. The proposed model is composed of four main approaches, which are ALT, ImgSibling, AllSubtree, and ParentSibling. The four main approaches exploit the used of Document Object Modeling by considering the contents that are the closest to the position of an image and the contents that are located around that position. Accordingly, a prototype of image search engine was developed and implemented as a tool for evaluating the extent of effectiveness of the proposed model. The underlying assumption is that if the proposed approaches select an appropriate content, the system will get a better

image index. In other words, the quality of the best image index is measured based on the effectiveness of the search results. The results of experiment revealed that the AllSubtree approach outperformed other approaches. The mean average precision was 0.74. Hence, the proposed model was a promising approach for the task of content selection. The model has the potential to be implemented as a process in an image indexing engine of any image search engine.

## 1. บทนำ

วิวัฒนาการของเทคโนโลยีสารสนเทศและเครือข่ายอินเทอร์เน็ต อำนวยความสะดวกให้กับคนในสังคมยุคใหม่เข้าถึงแหล่งข้อมูลต่าง ๆ ได้รวดเร็วมาก

ยิ่งขึ้น ดังจะเห็นได้จากข้อมูลการศึกษาผู้ใช้งานอินเทอร์เน็ตในประเทศไทย พ.ศ. 2560 [1] พบว่าคนที่มีช่วงอายุ 17 - 36 ปี เวลาบนอินเทอร์เน็ตต่อวันสูงที่สุด โดยใช้เวลา 7 ชั่วโมง 12 นาทีในวันทำงานหรือวันเรียนหนังสือ และ 7 ชั่วโมง 36 นาทีในวันหยุด ในขณะที่ภาพรวมเวลาที่คนไทยทุกช่วงอายุใช้อินเทอร์เน็ตต่อวัน อยู่ที่เวลา 6 ชั่วโมง 30 นาทีในวันทำงานหรือวันเรียนหนังสือ ในขณะที่ใช้เวลา 6 ชั่วโมง 48 นาทีในวันหยุด และกิจกรรมบนอินเทอร์เน็ตที่คนไทยทุกช่วงอายุนิยม 5 อันดับแรก ได้แก่ การใช้โซเชียลมีเดีย (86.9%) การค้นหาข้อมูล (86.5%) การรับส่งอีเมล (70.5%) การดูทีวีและฟังเพลงออนไลน์ (60.7%) และการซื้อสินค้าออนไลน์ (50.8%)



รูปที่ 1 ผลลัพธ์การค้นหารูปภาพบน Google.com โดยใช้คำว่า “หาดทรายแก้ว”

ตามรายงานของ Similarweb.com ที่ได้เก็บข้อมูลการเข้าเว็บไซต์ของคนไทยในเดือนมกราคม พ.ศ. 2560 พบว่า Google.co.th และ Google.com เป็นสองเว็บไซต์ที่คนไทยนิยมเข้าเป็นอันดับที่ 2 และอันดับที่ 4 ตามลำดับ ซึ่งสอดคล้องกับผลการศึกษาผู้ใช้งานอินเทอร์เน็ตในประเทศไทย [1] และเมื่อทดลองใช้ Google เพื่อค้นหารูปภาพ จะพบว่าผลลัพธ์ของการค้นคืนรูปภาพบางส่วนจะยังไม่ตรงกับคำค้นหาดังแสดงในรูปที่ 1 ซึ่งแสดงผลลัพธ์การค้นหารูปภาพบน โดยใช้คำค้นหาว่า “หาดทรายแก้ว” จะเห็นได้ว่าผลลัพธ์ที่ได้จากการค้นคืน มีรูปภาพส่วนหนึ่งเป็นรูปภาพห้องพัก และรูปกราฟิกของงานที่จะจัดขึ้นที่หาดทรายแก้ว เมื่อทำการวิเคราะห์ถึงปัญหาที่เกิดขึ้นพบว่า รูปภาพอื่นๆ ที่มาปรากฏเป็นส่วนหนึ่งของผลลัพธ์การค้นหา เกิดขึ้นจากการที่เนื้อหาของหน้าเว็บเพจที่มีรูปภาพนั้นๆ ปรากฏอยู่ มีคำว่า “หาดทรายแก้ว” เป็นส่วนหนึ่งของเนื้อหา ดังนั้นการที่เสิร์ชเอนจินนำเอาเนื้อหาทั้งหมดในเว็บเพจมาใช้สร้างเป็นดัชนีของรูปภาพจึงไม่น่าจะเป็นวิธีการที่เหมาะสม

จากงานวิจัยที่เกี่ยวข้องกับการค้นคืนรูปภาพในอดีตที่ผ่านมา จะพบว่ากระบวนจะถูกพัฒนาขึ้นในสองมุมมอง ได้แก่ 1) ระบบค้นคืนรูปภาพจากเนื้อหาของรูปภาพ (Content-Based Image Retrieval – CBIR) และ 2) ระบบค้นคืนรูปภาพจากข้อความ (Text Annotation Based Image Retrieval) นักวิจัยส่วนใหญ่ให้ความสนใจกับระบบค้นคืนรูปภาพจากเนื้อหารูปภาพเป็นส่วนใหญ่ ดังนั้นงานวิจัยตลอดระยะเวลาหลายสิบปีที่ผ่านมาจะมุ่งไปที่การพัฒนาเทคนิคในการสกัดคุณลักษณะ (Feature) ของรูปภาพเป็นหลัก [2] งานวิจัยในกลุ่มนี้จึงเน้นในเรื่องของการ

ใช้วิธีการประมวลผลภาพ ดังที่สรุปไว้ในงาน [3][4][5][6] ซึ่งหลาย ๆ งานมีการใช้การเก็บค่าข้อมูลคุณลักษณะของสี (Color Property) การวิเคราะห์รูปร่างของวัตถุ (Shape Analysis) รวมทั้งการวิเคราะห์พื้นผิว (Texture) [7] และนำเอาข้อมูลดังกล่าวมาช่วยในการตีความภาพและกำหนดดัชนีให้กับรูปภาพ อย่างไรก็ตามวิธีการที่เสนอในงานวิจัยในกลุ่มนี้ยังถูกนำมาประยุกต์ใช้ในเสิร์ชเอนจินไม่มากนัก เสิร์ชเอนจินที่ได้รับความนิยมในปัจจุบัน เช่น Google ยังคงใช้ประโยชน์จากงานวิจัยในกลุ่มระบบการค้นคืนรูปภาพจากข้อความเป็นหลัก [8] เนื่องจากเป็นแนวทางที่มีความรวดเร็วในการพัฒนา งานวิจัยในกลุ่มนี้ได้แก่ [9][10][11][12][13] ได้เสนอแนวคิดในการใช้ป้ายระบุข้อความ (Tag) เพื่อการสร้างเป็นตัวแทนของทรัพยากรเว็บ (Web Resources) และ [14] และ [15] ได้ทำการพิสูจน์แนวคิดดังกล่าวโดยเพื่อทำการพิสูจน์ว่าป้ายระบุสามารถนำไปใช้เป็นดัชนีแทนรูปภาพได้หรือไม่ ซึ่งผลการทดลองพบว่าผู้ค้นหารูปภาพส่วนใหญ่มักจะใช้คำค้นที่แตกต่างจากป้ายระบุข้อความ ทั้งนี้เนื่องมาจากลักษณะของป้ายระบุข้อความที่มีลักษณะเป็น Uncontrolled Vocabulary จึงส่งผลลัพธ์โดยตรงต่อการค้นหา

ดังนั้นในการพัฒนาเสิร์ชเอนจินเพื่อการค้นหารูปภาพ การใช้ข้อความที่อยู่บนหน้าเว็บจึงเป็นทางเลือกที่น่าสนใจและเป็นทางเลือกที่สะดวกและรวดเร็วที่สุด อย่างไรก็ตาม เนื้อหาของเว็บเพจทั้งหมดนั้น บางครั้งไม่เหมาะสมที่จะนำมาสร้างเป็นดัชนี ดังนั้นการหาสร้างตัวแบบที่มีความสามารถในการเลือกข้อความที่เหมาะสมที่จะนำไปสร้างเป็นดัชนีของรูปภาพจึงเป็นงานที่ท้าทาย

งานวิจัยชิ้นนี้นำเสนอตัวแบบสำหรับการคัดเลือกข้อความบนหน้าเว็บเพจ เพื่อการสร้างดัชนีของรูปภาพบนเว็บโดยใช้ตัวแบบของเอกสาร เพื่อปรับปรุงประสิทธิภาพของการค้นคืนรูปภาพ โดยผู้วิจัยคาดว่าตัวแบบที่พัฒนาขึ้นจะสามารถนำไปพัฒนาต่อให้เป็นกระบวนการอัตโนมัติ ที่จะทำให้การคัดเลือกข้อความบนเว็บเพจที่เหมาะสมที่สามารถนำไปใช้สร้างดัชนีของรูปภาพได้อย่างถูกต้อง

## 2. สมมติฐานและขอบเขตของงานวิจัย

### 2.1 สมมติฐาน

งานวิจัยนี้จะดำเนินการออกแบบและพัฒนาตัวแบบสำหรับการคัดเลือกข้อความบนหน้าเว็บเพจ เพื่อการสร้างดัชนีของรูปภาพบนเว็บโดยใช้ตัวแบบของเอกสาร และประเมินประสิทธิผลของตัวแบบที่พัฒนาขึ้น โดยเปรียบเทียบกับผลลัพธ์ของการค้นหารูปภาพ จากดัชนีรูปภาพที่ได้จากตัวแบบ เปรียบเทียบกับการผลลัพธ์ที่ได้จากดัชนีรูปภาพจากข้อความทั้งหมด

โดยมีสมมติฐานว่าตัวแบบสำหรับตัวแบบสำหรับการคัดเลือกข้อความบนหน้าเว็บเพจ เพื่อการสร้างดัชนีของรูปภาพบนเว็บโดยใช้ตัวแบบของเอกสาร มีความสามารถในการระบุเนื้อหาในเว็บเพจส่วนใดเป็นเนื้อหาที่เหมาะสมสำหรับการนำไปเป็นสร้างดัชนีที่เป็นตัวแทนของรูปภาพ ซึ่งเมื่อโปรแกรมการค้นหารูปภาพมีดัชนีที่เป็นตัวแทนของรูปภาพที่ดี ประสิทธิภาพของการค้นคืนรูปภาพก็จะดีขึ้นด้วย เมื่อเปรียบเทียบกับโปรแกรมการค้นหารูปภาพที่สร้างดัชนีรูปภาพด้วยจากข้อความทั้งหมด

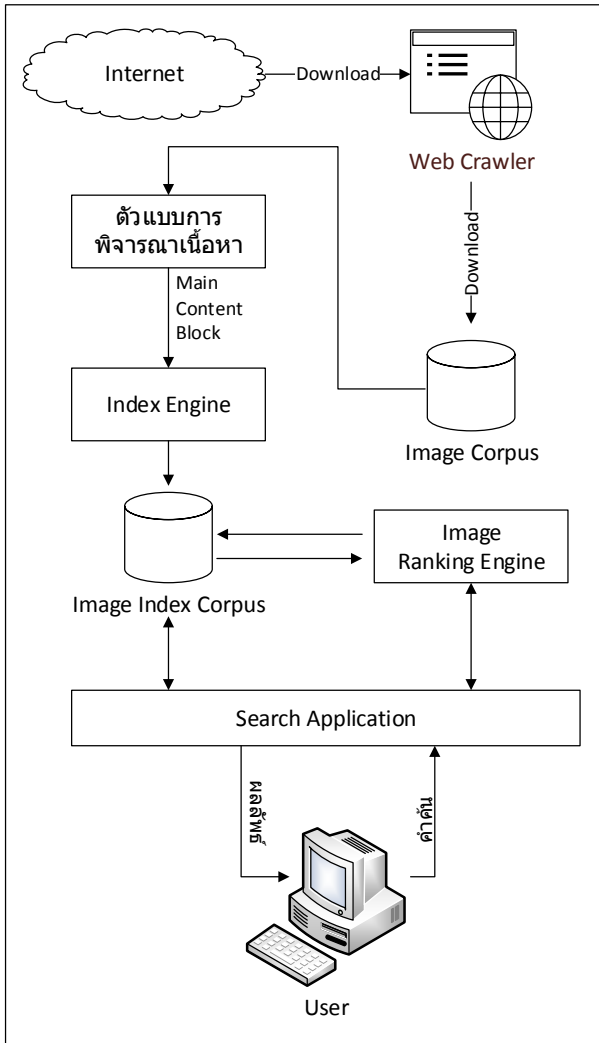
## 2.2 ขอบเขตของงานวิจัย

2.2.1 ตัวแบบที่พัฒนาขึ้นจะใช้วิธีการวิเคราะห์ตัวแบบของเอกสาร (Document Object Model) ร่วมกับตำแหน่งของรูปภาพที่ปรากฏอยู่บนเว็บเพจ เพื่อประเมินว่าเนื้อหาในเว็บเพจนั้น เกี่ยวข้องกับรูปภาพหรือไม่ ซึ่งจะส่งผลให้ได้ข้อความที่เหมาะสมสำหรับการนำไปสร้างดัชนีรูปภาพได้

2.2.2 การประเมินประสิทธิผลของตัวแบบ จะทำการเปรียบเทียบกับผลลัพธ์การค้นหารูปภาพจากดัชนีรูปภาพที่ได้จากตัวแบบ เปรียบเทียบกับผลลัพธ์การค้นหารูปภาพจากดัชนีที่ได้จากเนื้อหาทั้งหมดของเว็บเพจ โดยการประเมินผลลัพธ์ของการค้นหารูปภาพ จะใช้ผู้ร่วมทดลอง (Research Participants) จำนวนทั้งสิ้น 25 ท่าน

## 3. สถาปัตยกรรมระบบการค้นคืนรูปภาพและการออกแบบตัวแบบสำหรับการคัดเลือกข้อความ

เนื่องจากตัวแบบการคัดเลือกข้อความบนหน้าเว็บเพจ ที่นำเสนอในบทความนี้ จะถูกนำไปพัฒนาเป็นส่วนหนึ่งของการทำงานของเว็บครอว์เลอร์ (Web Crawler) และ เครื่องสร้างดัชนี (Index Engine) โดยมีหน้าที่หลักในการระบุส่วนของเนื้อหาหลัก (Main Content Block) ที่น่าจะนำมาใช้สำหรับการสร้างดัชนีรูปภาพ ดังนั้นในส่วนนี้จะทำการอธิบายภาพรวมของสถาปัตยกรรมระบบการค้นคืนรูปภาพ และส่วนประกอบหลักของระบบ (Main System Components) จะประกอบไปด้วย 1) เว็บครอว์เลอร์ (Web Crawler) ซึ่งทำหน้าที่ในการดาวน์โหลดเนื้อหา (Content) และรูปภาพจากเว็บเพจต่างๆ นำมาเก็บไว้ที่



รูปที่ 2 แสดงสถาปัตยกรรมและส่วนประกอบหลักของระบบค้นคืนรูปภาพ

2) คลังรูปภาพ (Image Corpus) ซึ่งในคลังรูปภาพจะมีการเก็บเนื้อหาบนเว็บเพจ (Content) ที่จะถูกนำมาสร้างเป็นดัชนี โดยเนื้อหาดังกล่าวจะถูกพิจารณาด้วย 3) ตัวแบบการพิจารณาเนื้อหาเว็บเพจ ซึ่งจะทำหน้าที่ระบุส่วนของเนื้อหาหลัก (Main Content Block) เพื่อส่งต่อไปให้ 4) เครื่องสร้างดัชนี (Index Engine) โดยจะทำหน้าที่นำเอาเนื้อหาไปสร้างดัชนี เพื่อเก็บไว้ใน 5) คลังดัชนีรูปภาพ (Image Index Corpus) ซึ่งจะเก็บดัชนีของรูปภาพ โดยจะถูกใช้งานโดย 6) แอปพลิเคชัน

ขั้นการค้นหา และ 7) เครื่องเรียงลำดับผลการค้นคืน (Ranking Engine ดังในรูปที่ 2 แสดงสถาปัตยกรรมและส่วนประกอบหลักของระบบการค้นคืนรูปภาพ โดยรูปดังกล่าวปรับปรุงมาจาก [16]

```

<div class="panel-grid-cell" id="pgc-146-1-1" >
<div class="so-panel widget widget_edupress edupress panel-
first-child panel-last-child" id="panel-146-1-1-0" data-index="3">
<div style="margin-bottom: 30px;" class="panel-widget-style" >
<div class="textwidget">
<div class="row">
<div class="col-md-4">
<div class="team-member-img">

</div>
</div>
</div>
<div class="col-md-8">
<div class="single-team-member">
<div class="team-member-name">
<p>ผู้ช่วยศาสตราจารย์ ธีรนาถ พงษ์</p>
</div>
<div class="team-member-link">
<a href="#">
<i class="fa fa-facebook"></i>
</a>
<a href="#">
<i class="fa fa-twitter"></i>
</a>
<a href="#">
<i class="fa fa-linkedin"></i>
</a>
</div>
</div>
</div>
</div>
</div>
</div>
</div>

```

รูปที่ 3 แสดง HTML Code

จาก <http://ant.dpu.ac.th/index.php/lecturers/>

เนื่องจากงานวิจัยชิ้นนี้ จะเน้นที่การคัดเลือกข้อความที่อยู่บนเว็บเพจ เพื่อนำมาใช้เป็นดัชนีตัวแทนรูปภาพ โดยจะใช้แนวทางการวิเคราะห์ตัวแบบเอกสาร ดังนั้นในงานวิจัยชิ้นนี้ จึงเสนอแนวทางในการในการคัดเลือกข้อความที่จะนำมาใช้เพื่อสร้างดัชนีรูปภาพ ดังนี้

แนวทางที่ 1 ใช้ค่าของข้อมูลใน alt Attribute ของ <img> Tag ดังแสดงในรูปที่ 3 แสดง HTML ในส่วนของ alt จาก <http://ant.dpu.ac.th/index.php/lecturers/> โดยปกติแล้ว alt หรือ Alternate Text จะใช้สำหรับการใส่ข้อความที่อธิบายรูปภาพ ซึ่งจะใช้โดย

เครื่องมืออ่านหน้าจอ (Screen Reader) สำหรับผู้ใช้เว็บที่มีปัญหาทางสายตา หรือจะใช้แสดงในกรณีที่รูปภาพที่กำหนดไว้ใน `<img>` ไม่สามารถแสดงผลได้ ดังนั้นข้อความที่ระบุไว้ใน `alt` จึงเป็นข้อความที่น่าสนใจที่จะนำมาใช้สำหรับการสร้างดัชนีตัวแทนรูปภาพ

แนวทางที่ 2 จะพิจารณาจากลักษณะของการวางข้อความเพื่ออธิบายรูปภาพ มักจะอยู่ในรูปแบบวางอยู่ข้างๆ กัน ดังรูปที่ 4 แสดง HTML การจัดวางรูปภาพ และข้อความอธิบายรูปภาพ จาก <http://ant.dpu.ac.th/> และรูปที่ 5 แสดงภาพจำลองตัวแบบเอกสารจาก HTML ในรูปที่ 4

```

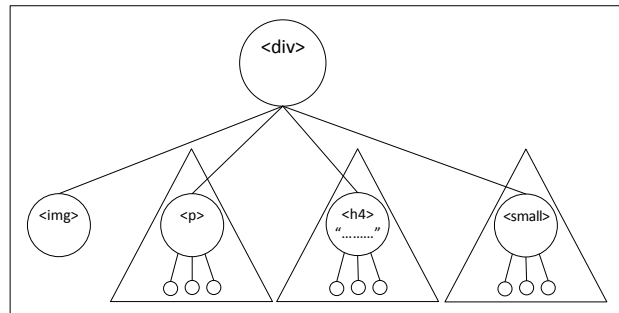
<div class="col-md-6 col-sm-6">
  <div class="testibox">
    
    <p style="color: #676767;">สร้างสรรคด้วยจินตนาการ ผสมผสานศาสตร์แห่งศิลป์และเทคโนโลยี</p>
    <h4 style="color: #27293d;">ผู้ช่วยศาสตราจารย์ วิฑูญ นรินทรสาพงศ์</h4>
    <small style="color: #676767;">รองคณบดีฝ่ายวิชาการ at <a href="#" style="color: #676767;">ANT</a></small>
  </div>
</div>

```

**<img> และ <h4> เป็น Sibling อยู่ภายใน <div> ซึ่งเป็น Parent Node**

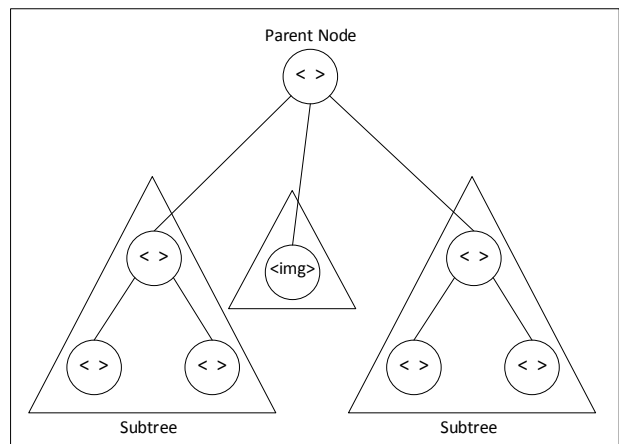
รูปที่ 4 แสดง HTML การจัดวางรูปภาพและข้อความอธิบายรูปภาพ

ในรูปที่ 4 และ 5 จะเห็นได้ว่า `<h4>` มีข้อความอธิบายรูปภาพที่กำหนดไว้ `<img>` ซึ่งเมื่อพิจารณาการจัดวางในรูปแบบของตัวแบบเอกสารแล้ว จะพบว่า `<h4>` และ `<img>` เป็น Sibling Node โดยมี `<div>` เป็น Parent Node ดังนั้นจึงเรียกแนวทางนี้ว่า "ImgSibling"



รูปที่ 5 แสดงภาพจำลองตัวแบบเอกสารจาก HTML ในรูปที่ 4

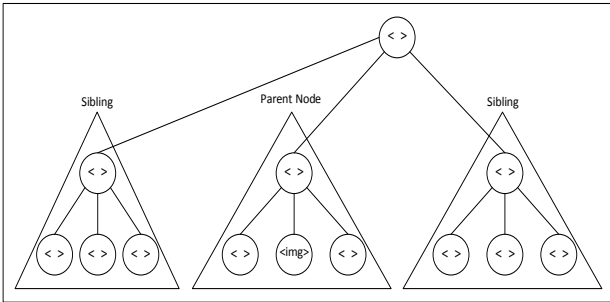
แนวทางที่ 3 จะใช้ข้อมูลของทั้ง Subtree ของ Parent Node ที่มี `<img>` เป็น Child Node ดังแสดงไว้ในรูปที่ 6 โดยแนวทางนี้จะพิจารณาจากวิธีการวาง Layout ของเนื้อหาและรูปภาพ เมื่อเปรียบเทียบกับแนวทางที่ 2 แนวทางนี้เป็นการให้ความสำคัญกับเนื้อหาที่อยู่บริเวณใกล้เคียงกับรูปภาพในด้านลึก (Depth) ซึ่งเป็นการเพิ่มปริมาณข้อความที่จะนำมาใช้สร้างดัชนีรูปภาพ โดยจะเรียกแนวทางนี้ว่า "AllSubtree"



รูปที่ 6 แสดงภาพจำลองตัวแบบ AllSubtree

สำหรับแนวทางสุดท้าย จะเป็นการให้ความสำคัญกับเนื้อหาที่อยู่บริเวณรอบข้างรูปภาพในด้านกว้าง (Breadth) จะเรียกแนวทางนี้ว่า "ParentSibling" โดยแนวทางนี้จะพิจารณาเนื้อหา

ของ Subtree ของ Sibling Node ของ Parent Node ที่มี <img> อยู่ ดังที่แสดงในรูปที่ 7 จะเห็นได้ว่าแนวทางนี้เป็นการพิจารณาข้อความที่อยู่ห่างจากรูปภาพระดับหนึ่งแต่ยังมีความน่าจะเป็นที่จะเป็นข้อความที่ใช้อธิบายรูปภาพได้



รูปที่ 7 แสดงภาพจำลองตัวแบบ ParentSibling

จะเห็นได้ว่าทั้ง 4 แนวทางที่เสนองานในบทความวิจัยฉบับนี้ ใช้หลักการคัดเลือกข้อความจากตำแหน่งที่ใกล้กับตำแหน่งรูปภาพมากที่สุด นั่นก็คือค่าของ Attribute ภายใน <img> และค่อยๆ ขยับตำแหน่งออกไปในส่วนที่ไกลออกไปแต่ยังมีความน่าจะเป็นที่จะเป็นข้อความอธิบายรูปภาพได้ โดยใช้การวิเคราะห์ตัวแบบเอกสาร (Document Object Modeling Analysis)

#### 4. การทดลองและการประเมินผลตัวแบบ

##### 4.1 การสร้างคลังรูปภาพและการเตรียมดัชนีรูปภาพ

ในการประเมินประสิทธิผลของตัวแบบ คลังรูปภาพ (Image Corpus) จึงได้ถูกสร้างขึ้น โดยการพัฒนาและส่งเว็บครอว์เลอร์ (Web Crawler) ไปทำการดาวน์โหลดรูปภาพจากเว็บเพจ โดยกำหนดเว็บเริ่มต้น (Seed Web) ซึ่งเป็นเว็บที่เกี่ยวข้องกับการท่องเที่ยว จำนวน 15 โดเมน สาเหตุที่เลือกเว็บเริ่มต้นมีเนื้อหาเกี่ยวข้องกับการท่องเที่ยว เพราะเหตุว่า

ลักษณะของเนื้อหา มักจะประกอบไปด้วยรูปภาพและคำอธิบายภาพ จึงเป็นแหล่งข้อมูลที่ดีที่จะใช้ประเมินประสิทธิผลของตัวแบบ

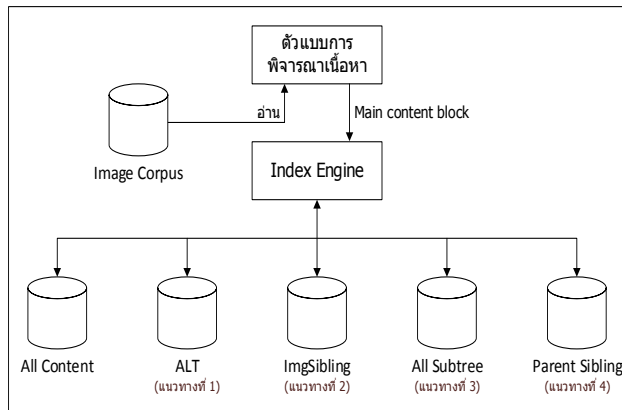
เพื่อให้สามารถรวบรวมคลังรูปภาพได้ในปริมาณเหมาะสม เว็บครอว์เลอร์ถูกส่งออกไปทำการเก็บข้อมูลในช่วงเดือนมีนาคม ถึงเดือนสิงหาคม 2559 โดยได้เข้าไปดาวน์โหลดเว็บเพจมากกว่า 1 หมื่นเว็บเพจ และได้รูปภาพจำนวน 94,306 รูป พร้อมกับเนื้อหาของหน้าเว็บเพจดังแสดงในรูปที่ 8 จากนั้นข้อมูลทั้งหมดจะถูกอ่านโดยตัวแบบการพิจารณาเนื้อหา และทำการระบุ Main Content Block ที่จะทำมาเป็นข้อมูลหลักสำหรับใช้สร้างดัชนีรูปภาพ โดยใช้แนวทางในการคัดเลือกเนื้อหาทั้ง 4 แนวทางดังรายละเอียดที่กล่าวมาในหัวข้อที่ผ่านมา

Table	Action	Records
image	[Icons]	94,306
seed	[Icons]	10,102
2 table(s) Sum		104,408

รูปที่ 8 แสดงจำนวนรูปภาพและเว็บเพจที่เก็บอยู่ในฐานข้อมูลของระบบ

ดังนั้นในการประเมินประสิทธิผลของตัวแบบทำการเปรียบเทียบผลลัพธ์การค้นคืนรูปภาพที่จากดัชนีรูปภาพ 5 ดัชนี ได้แก่ ดัชนีรูปภาพที่ได้จากนำเอาเนื้อหาทั้งหน้าเว็บเพจมาพิจารณา ซึ่งเป็นแนวทางมาตรฐานที่ระบบค้นคืนรูปภาพ เช่น Google ใช้อยู่ในปัจจุบัน ดัชนีรูปภาพที่สร้างจากแนวทาง ALT, แนวทาง ImgSibling, แนวทาง AllSubtree, และ

แนวทาง ParentSibling ดังที่แสดงในรูปที่ 9 และรูปที่ 10



รูปที่ 9 แสดงการเตรียมดัชนีรูปภาพสำหรับการประเมินประสิทธิภาพผลตัวแบบ

การสร้างดัชนีแต่ละแบบจะใช้แบบจำลองปริภูมิเวกเตอร์ (Vector Space Model) โดยกำหนดค่า

น้ำหนักในรูป TF/IDF (Term Frequency /Inverse Document Frequency) ดังแสดงในสมการที่ 1

$$w_{i,j} = \frac{f_{i,j}}{\sum_{k=1}^m f_{k,j}} \times \log \frac{N}{n_i} \quad (1)$$

โดยที่

$w_{i,j}$  หมายถึงค่าน้ำหนักของ Term ที่นำมาสร้างเป็นดัชนีของรูปภาพ

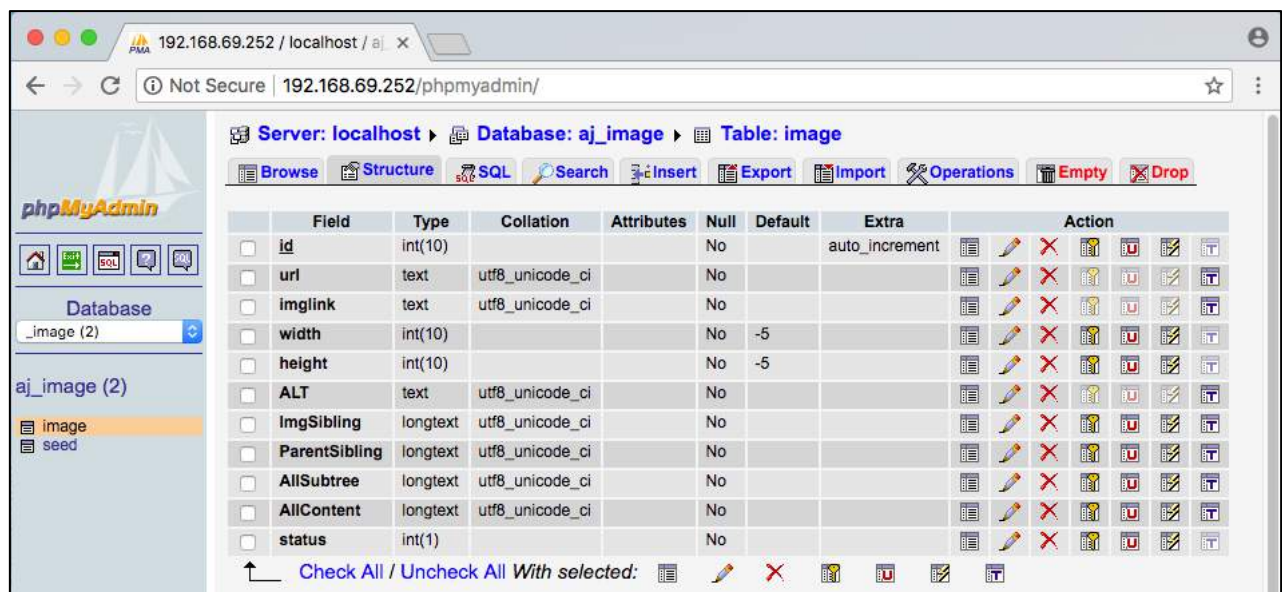
$f_{i,j}$  หมายถึงความถี่ของ Term i ในเอกสารที่ j

$m$  หมายถึงจำนวน Unique Term ในเอกสาร j

$n_i$  หมายถึงจำนวนเอกสารที่มี Term i ปรากฏอยู่

$N$  หมายถึงจำนวนเอกสารทั้งหมดในคลังเอกสาร

และเอกสารในที่นี้หมายถึง รูปภาพและข้อความที่เลือกโดยตัวแบบ



รูปที่ 10 แสดงโครงสร้างฐานข้อมูล Main Content Block



## 4.2 การออกแบบการทดลองและการวัดประสิทธิผล ตัวแบบสำหรับการคัดเลือกข้อความ

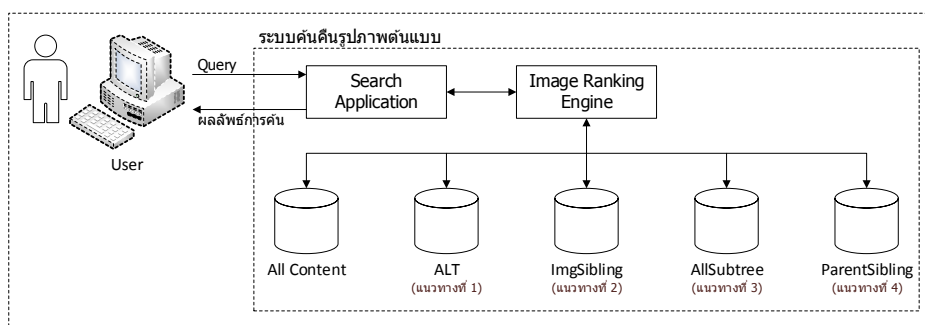
เนื่องจากตัวแบบที่เสนอในงานวิจัยชิ้นนี้ เป็นตัวแบบสำหรับการคัดเลือกข้อความที่เหมาะสมที่จะนำไปใช้ทำดัชนีรูปภาพเท่านั้น ในการประเมินประสิทธิผลการทำงานของตัวแบบจึงใช้สมมติฐานที่ว่า “ถ้าตัวแบบคัดเลือกข้อความได้เหมาะสม จะทำให้ได้ดัชนีรูปภาพที่ดี และดัชนีรูปภาพที่ดี จะส่งผลทำให้ผลลัพธ์การค้นคืนรูปภาพดีด้วย”

ดังนั้นการประเมินประสิทธิผลของตัวแบบสำหรับการคัดเลือกเนื้อหาเว็บเพจที่เสนอในงานวิจัยชิ้นนี้ จะจัดในรูปแบบของการทำการทดลอง (Experiment) โดยเชิญ ผู้ร่วมทดลอง (Research Participants) จำนวนทั้งสิ้น 25 คน โดยเป็นผู้ที่สนใจในเรื่องของการค้นหาสถานที่ท่องเที่ยวและมีความคุ้นเคยกับการใช้งานเสิร์ชเอนจิน และในการทดลองนั้น จะมอบหมายให้ผู้ร่วมทดลองทำการค้นหารูปภาพสถานที่ท่องเที่ยวตามความสนใจของผู้ร่วมทดลองจำนวน 4 แห่ง โดยสามารถกำหนดคำค้น

(Query) ได้ตามต้องการ โดยจะทำให้ได้ผลลัพธ์การสืบค้นทั้งสิ้น 100 ชุด เพื่อจะทำไปประเมินผลการค้นคืน โดยผู้เข้าร่วมทดลองจะทำการประเมินผลลัพธ์ที่ได้รับมาจากระบบค้นคืนรูปภาพต้นแบบ โดยระบุว่าเป็นรูปภาพที่ถูกต้องตรงกับสิ่งที่ค้นหาหรือไม่ โดยจะระบุ ‘1’ ซึ่งหมายถึงตรงกับคำสืบค้น และ ‘0’ ซึ่งหมายถึงไม่ตรงกับคำสืบค้น

ระบบค้นคืนรูปภาพต้นแบบถูกพัฒนาขึ้นเพื่อใช้สำหรับการทดลองในครั้งนี้ โดยในแต่ละครั้งที่ผู้ร่วมทดลองทำการค้นหารูปภาพ ระบบจะนำคำค้นไปทำการเปรียบเทียบกับดัชนีรูปภาพทั้งหมด 5 ครั้งโดยใช้ดัชนี AllContent, ดัชนี ALT, ดัชนี ImgSibling, ดัชนี AllSubtree และดัชนี ParentSibling ดังแสดงในรูปที่ 10 และในการคัดเลือกรูปภาพมาเป็นผลลัพธ์การค้นคืนจะใช้การวัดความเหมือนแบบโคไซน์ (Cosine Similarity) ดังแสดงในสมการที่ 2

$$Sim(Q, D) = \frac{\sum_{i=1}^n Q_i D_i}{\sqrt{\sum_{i=1}^n Q_i^2} \sqrt{\sum_{i=1}^n D_i^2}} \quad (2)$$



รูปที่ 11 แสดงส่วนประกอบของการค้นคืนรูปภาพ

โดยระบบจะทำการคัดเลือก 15 ลำดับแรก ที่มีคะแนนความเหมือนแบบโคไซน์ของแต่ละแนวทาง และจะนำผลลัพธ์การค้นคืนดังกล่าวรวมเข้าด้วยกัน

เพื่อเป็นผลลัพธ์การค้นคืนชุดเดียว ซึ่งเป็นการอำนวยความสะดวกให้กับผู้ร่วมทดลอง ในการประเมินความเกี่ยวพัน (Relevancy Rating) นอกจากนั้นแล้ว

ผลลัพธ์การค้นคืนรูปภาพที่ได้ จะนำเสนอให้กับผู้ร่วมทดลองแบบสุ่ม (Randomized Order) และผู้ควบคุมการทำการทดลองจะแจ้งให้กับผู้ร่วมทดลองให้ทราบว่าผลลัพธ์จะแสดงแบบสุ่ม โดยวิธีการดังกล่าวจะช่วยลดความลำเอียง (Bias) ที่อาจจะเกิดขึ้นจากลำดับการนำเสนอผลลัพธ์การค้นคืน

ผลลัพธ์การค้นคืนแต่ละแนวทางจะถูกบันทึกไว้พร้อมกับผลการประเมินรูปภาพ และจะนำเอาผลลัพธ์ดังกล่าวมาคำนวณ ค่าความแม่นยำ (Precision) โดยจะนำผลการประเมินของผู้เข้าร่วมทดลองในทุกคำค้นมาคำนวณหาค่าความแม่นยำ และนำดังกล่าวของทุกคำค้นของผู้เข้าร่วมทดลองมาหาค่าเฉลี่ยของความแม่นยำของแต่ละแนวทาง เพื่อนำมาเปรียบเทียบประสิทธิผลของแต่ละแนวทางจากดัชนีในการทำการทดลอง

ผลลัพธ์การค้นคืนแต่ละแนวทางจะถูกบันทึกไว้พร้อมกับผลการประเมินรูปภาพ และจะนำเอาผลลัพธ์ดังกล่าวมาคำนวณ ค่าความแม่นยำ (Precision) โดยจะนำผลการประเมินของผู้เข้าร่วมทดลองในทุกคำค้นมาคำนวณหาค่าความแม่นยำ และนำดังกล่าวของทุกคำค้นของผู้เข้าร่วมทดลองมาหาค่าเฉลี่ยของความแม่นยำของแต่ละแนวทาง เพื่อนำมาเปรียบเทียบประสิทธิผลของแต่ละแนวทาง

## 5. ผลการประเมินตัวแบบ

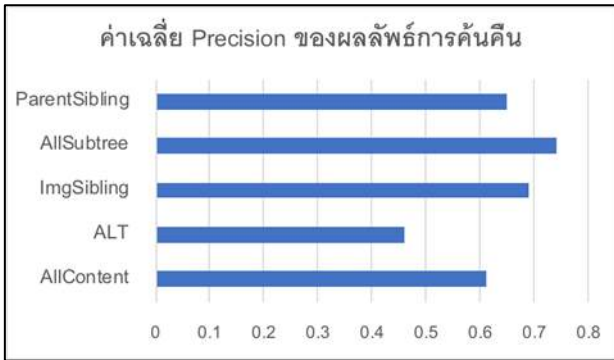
จากผลการทดลองพบว่าผลลัพธ์การค้นคืนที่ได้จากดัชนี AllSubtree มีค่าความแม่นยำเฉลี่ยสูงที่สุด ได้แก่ 0.74 รองลงมาเป็นของดัชนี ImgSibling ให้ค่าความแม่นยำเฉลี่ย 0.69 และตามด้วยดัชนี ParentSibling ให้ค่าความแม่นยำเฉลี่ย 0.65 ในขณะที่

ที่ดัชนี AllContent และ ALT ให้ค่าความแม่นยำเฉลี่ย 0.61 และ 0.46 ตามลำดับ ดังแสดงในตารางที่ 1 และรูปที่ 11 แสดงกราฟเปรียบเทียบค่าเฉลี่ย Precision ของผลลัพธ์การค้นคืน

เมื่อพิจารณาในรายละเอียดของตัวแบบในการคัดเลือกข้อความแต่ละแบบที่นำเสนอในงานวิจัยชิ้นนี้ จะพบว่าตัวแบบ AllSubtree มีความสามารถในการคัดเลือกข้อความที่เป็นตัวแทนของรูปภาพได้ดีที่สุด เนื่องจากมีการพิจารณาการวาง Layout ของข้อความและรูปภาพ และให้ความสำคัญกับข้อความที่อยู่บริเวณรอบๆ รูปภาพ รวมทั้งยังพิจารณาข้อความอื่นๆ เพิ่มเติมในแนวลึก (Depth) ของแต่ละ Subtree ด้วย ในขณะที่ตัวแบบ ImgSibling มีการพิจารณาข้อความที่อยู่ข้างๆ รูปภาพ ในแนวกว้าง (Breadth) แต่ไม่ได้พิจารณาลงลึกในแต่ละ Subtree ย่อย จึงส่งผลให้อาจจะพลาดข้อความที่สำคัญที่เหมาะสมที่จะนำไปสร้างดัชนีตัวแทนรูปภาพ ในทางกลับกัน ตัวแบบ ParentSibling จะพิจารณาข้อความที่อยู่รอบๆ รูปภาพทั้งในแนวกว้างและลึก จึงส่งผลทำให้ดัชนีที่เป็นตัวแทนรูปภาพมีความถูกต้องน้อยลง เนื่องจากมีการนำเอาข้อความที่ไม่เกี่ยวข้องกับรูปภาพมาสร้างเป็นดัชนี

ตารางที่ 1 ตารางค่าเฉลี่ย Precision ของผลลัพธ์การค้นคืน

แนวทางการคัดเลือกข้อความมาใช้สร้างดัชนี	ค่าเฉลี่ย Precision ของผลลัพธ์การค้นคืน
ParentSibling	0.65
AllSubtree	0.74
ImgSibling	0.69
ALT	0.46
AllContent	0.61



รูปที่ 11 แสดงกราฟเปรียบเทียบค่าเฉลี่ยของผลลัพธ์การค้นหา  
คือรูปภาพ

สำหรับตัวแบบ ALT นั้น ถ้าพิจารณาจากแนวทางแล้ว น่าจะเป็นวิธีการที่ได้ดัชนีที่เป็นตัวแทนรูปภาพที่ดีที่สุด เนื่องจากเป็นการนำเอาข้อความที่เป็นคำอธิบายรูปภาพมาใช้สำหรับการสร้างดัชนี แต่ผลการทดลองกลับให้ค่าความแม่นยำเฉลี่ยต่ำที่สุด ทั้งนี้เป็นเพราะว่าในหลายๆ เว็บเพจ มักจะไม่กำหนดค่าข้อมูลให้กับ ALT Attribute หรือ กำหนดค่าข้อความที่เป็นคำอธิบายรูปภาพกว้างๆ ส่งผลทำให้ดัชนีที่ถูกสร้างขึ้นขาดความสมบูรณ์และเป็นตัวแทนที่ไม่ดีสำหรับรูปภาพ

## 6. บทสรุปและข้อเสนอแนะ

บทความวิจัยชิ้นนี้ นำเสนอตัวแบบสำหรับการคัดเลือกข้อความบนหน้าเว็บเพจสำหรับนำไปสร้างเป็นดัชนีรูปภาพโดยการใช้อย่างมีประสิทธิภาพเป็นแนวทางในการพิจารณา โดยมีการนำเสนอแนวทางในการคัดเลือกทั้งหมด 4 แนวทางได้แก่ ALT, ImgSibling, AllSubtree และ ParentSibling ทั้ง 4 แนวทางดังกล่าวเป็นการใช้หลักการวิเคราะห์ที่ตัวแบบเอกสารโดยพิจารณาข้อความที่ใกล้กับตำแหน่งรูปภาพมากที่สุด และค่อยๆ ขยับตำแหน่งออกไปใน

ส่วนที่ไกลออกไป ดังนั้นในการประเมินประสิทธิผลของตัวแบบจึงจัดในรูปแบบการทดลอง โดยมีผู้เข้าร่วมทดลองทั้งหมด 25 คน โดยแต่ละคนได้รับการมอบหมายให้ทำการค้นหารูปภาพสถานที่ท่องเที่ยวตามความสนใจจำนวน 4 แห่ง และทำการประเมินผลลัพธ์การค้นหาผ่านระบบค้นหารูปภาพต้นแบบ ผลลัพธ์การค้นหาและผลการประเมินรูปภาพจะถูกบันทึกไว้ และนำมาคำนวณค่าความแม่นยำจากผลการทดลองพบว่า ผลลัพธ์การค้นหาที่ได้จากแนวทาง AllSubtree มีค่าความแม่นยำสูงที่สุด ซึ่งได้แก่ 0.74 ในขณะที่แนวทาง ImgSibling และ ParentSibling ให้ค่าความแม่นยำเฉลี่ย 0.69 และ 0.65 ตามลำดับ แนวทางที่เสนอในงานวิจัยชิ้นนี้แม้ว่าจะเป็นการเสนอตัวแบบอย่างง่ายๆ แต่สามารถนำไปพัฒนาให้เป็นกระบวนการทำงานส่วนหนึ่งของเครื่องสร้างดัชนีรูปภาพ (Image Index Engine) ของระบบค้นหารูปภาพ ดังนั้นการปรับปรุงตัวแบบโดยใช้แนวทางอื่นๆ รวมทั้งการปรับปรุงแนวทางที่เสนอในบทความนี้ เป็นสิ่งที่ควรดำเนินการต่อ เพื่อเพิ่มคุณภาพดัชนีรูปภาพของระบบค้นหารูปภาพในอนาคต

## กิตติกรรมประกาศ

งานวิจัยชิ้นนี้ ได้รับทุนสนับสนุนการทำวิจัยจากมหาวิทยาลัยธุรกิจบัณฑิต และขอขอบคุณคณาจารย์และนักศึกษาในสาขาวิชาวิศวกรรมเว็บและการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา ที่สละเวลามาเป็นผู้เชี่ยวชาญในการประเมินเนื้อหาเว็บเพจเพื่อสร้างชุดคำตอบที่ถูกต้อง ซึ่งถูกนำมาใช้สำหรับการประเมินตัวแบบ

## เอกสารอ้างอิง

- [1] รายงานผลการสำรวจพฤติกรรมผู้ใช้อินเทอร์เน็ตของประเทศไทย ปี 2560 สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (องค์การมหาชน) กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม (<https://www.etcha.or.th/content/thailand-internet-user-profile-2017.html>)
- [2] K. Seetharaman and W.T. Chembian, "Statistical Distribution-based Color Image Retrieval", Proceedings of the International Conference on Graphics and Signal Processing, 2017, pp. 6-9
- [3] A. Aker, X. Fan, M. Sanderson, and R. Gaizauskas., "Investigating summarization techniques for geo-tagged image indexing", Proceedings of the 34<sup>th</sup> European Conference on Advances in Information Retrieval, 2012, pp.472-475
- [4] A.P. Bhagat and M. Atique, "Medical Image Retrieval, Indexing and Enhancement Techniques: A Survey", Proceedings of the 2011 International Conference on Communication, Computing & Security, 2011, pp.387-390
- [5] T. Chiueh, "Content-based Image Indexing", Proceedings of the 20th International Conference on Very Large Data Bases, 1994, pp. 582-593
- [6] E.A. Ei-Kwae and M. Kabuka, "Efficient Content-based Indexing of Large Databases", ACM Transaction on Information Systems, 18(2), pp. 171-210
- [7] R. Jain, P.M. Sudha, S.K. Pramod, and C.V. Jawahar, "An Indexing Approach for Speeding-up Image Classification", Proceedings of the 7th Indian Conference on Computer Vision, Graphics and Image Processing, 2010, pp. 290-297
- [8] M. Hu, Y. Yang, F. Shen, L. Zhang, H.T. Shen, and X. Li, "Robust Web Image Annotation via Exploring Multi-Facet and Structural Knowledge", IEEE Transactions on Image Processing, 26(10), 2017, pp. 4471-4884
- [9] R. Lukac and K.N. Plataniotis, "A New Approach to Camera Image Indexing", Proceedings of the 11<sup>th</sup> International Conference on Computer Analysis of Images and Patterns", 2005, pp. 137-144
- [10] C. Marlow, M. Naaman, D. Boyd, and M. Davis, "Tagging paper, taxonomy, Flickr, Academic Article", Proceedings of the 17<sup>th</sup> International Conference on Hypertext and Hypermedia, 2006, pp. 31-40
- [11] C. Luo, B. Ni, S. Yan, and M. Wang, "Image Classification by Selective Regularized Subspace Learning", IEEE Transactions on Multimedia, 18(1), 2016, pp. 40-50
- [12] X. Li, B. Shen, B.D. Liu, and Y.J. Zhang, "A Locality Sensitive Low Rank Model for

- Image Tag Completion”, IEEE Transactions on Multimedia, 18(3), 2016, pp. 474-483
- [13] M. Wang, W. Li, D. Liu, B. Ni, J. Shen, and S. Yan, “Facilitating Image Search with a Scalable and Compact Semantic Mapping”, IEEE Transactions on Cybernetics, 45(8), 2015, pp. 1561-1574
- [14] Y.M. Kim, “Social Tag in Text and Image Search”, Proceedings of the 3rd Symposium on Information Interaction in Context, 2010, pp. 353-356
- [15] Y. Yang, Z.J. Zha, Y. Gao, X. Zhu, and T.S. Chua, “Exploiting Web Images for Semantic Video Indexing via Robust Sample-specific Loss”, IEEE Transactions on Multimedia, 16(6), 2014, pp. 1677-1689
- [16] วรสิทธิ์ ชูชัยวัฒนา, “การปรับปรุงประสิทธิภาพของระบบค้นคืนสารสนเทศและโปรแกรมการค้นหาค้นหา: แนวคิดและเทคนิค”, วารสารวิชาการสมาคมสถาบันอุดมศึกษาเอกชนแห่งประเทศไทย ฉบับวิทยาศาสตร์และเทคโนโลยี ปีที่ 3 ฉบับที่ 1 มกราคม - มิถุนายน พ.ศ. 2557 หน้า 73-83.