รายงานวิจัยฉบับสมบูรณ์

การศึกษาการใช้ปัญญาประดิษฐ์เพื่อพัฒนาเมทาดาทาภาพถ่ายดิจิทัล

A study of using artificial intelligence to enhance digital image metadata

ผู้ช่วยศาสตราจารย์ ดร. อัครา ธรรมมาสถิตย์กุล

โครงการวิจัยประเภทงบประมาณเงินรายได้มหาวิทยาลัย
เงินรายได้ส่วนงาน มหาวิทยาลัยบูรพา
ประจำปีงบประมาณ พ.ศ. 2565
คณะมนุษยศาสตร์และสังคมศาสตร์
มหาวิทยาลัยบูรพา
2566

สัญญาเลขที่ HUSO๐๑/๒๕๖๕

รายงานวิจัยฉบับสมบูรณ์

การศึกษาการใช้ปัญญาประดิษฐ์เพื่อพัฒนาเมทาดาทาภาพถ่ายดิจิทัล

A study of using artificial intelligence to enhance digital image metadata

ผู้ช่วยศาสตราจารย์ ดร. อัครา ธรรมมาสถิตย์กุล
คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยบูรพา

ประกาศคุณูปการ

งานวิจัยเรื่องการศึกษาการใช้ปัญญาประดิษฐ์เพื่อพัฒนาเมทาดาทาภาพถ่ายดิจิทัล จัดทำขึ้นเพื่อสำรวจเครื่องมือในการรู้จำวัตถุในภาพถ่ายที่ใช้ปัญญาประดิษฐ์เพื่อกำหนดเมทาดาทา และเพื่อวิเคราะห์และประเมินประสิทธิภาพเครื่องมือในการรู้จำภาพที่เหมาะสมกับภาพถ่ายแต่ละประเภท โดยได้รับทุนสนับสนุนการวิจัยจากคณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยบูรพา จึงขอขอบคุณไว้ ณ ที่นี้

ขอขอบคุณเจ้าของภาพถ่ายที่นำภาพมาเผยแพร่แบบไม่มีลิขสิทธิ์ หรือยกให้เป็นสาธารณะประโยชน์ ที่นำมาใช้ในงานวิจัยนี้ทุกท่าน ขอขอบคุณคณาจารย์และเจ้าหน้าที่ภาควิชาสารสนเทศศึกษาทุกท่าน ที่กรุณาให้คำปรึกษา ช่วยเหลือ และชี้แนะแนวทางในการทำวิจัย

ขอขอบพระคุณผู้ทรงคุณวุฒิ ผู้ช่วยศาสตราจารย์ ดร. ศรีหทัย เวลล์ส คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยบูรพา ผู้ช่วยศาสตราจารย์ ดร.ศิริกานต์ ชูเชิด สำนักวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยแม่ฟ้าหลวง ผู้ช่วยศาสตราจารย์ ดร. ณิชนันทน์ กิตติพัฒนบวร สำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์ รองศาสตราจารย์ ดร. อัณณ์นุพันธ์ รอดทุกข์ ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยรามคำแหง และผู้เกี่ยวข้องทุกท่านที่กรุณาตรวจสอบ ปรับปรุง แก้ไขข้อบกพร่อง และให้คำแนะนำในการทำรายงานการวิจัยให้ถูกต้องสมบูรณ์ยิ่งขึ้น

สุดท้ายนี้ ขอขอบพระคุณ ครู อาจารย์ ทุกท่านที่ประสิทธิ์ประสาทวิชาความรู้ในการทำวิจัยนี้ให้สำเร็จ และขอบคุณสมาชิกในครอบครัวที่ส่งเสริมกำลังใจตลอดมา

ผู้วิจัย

หัวข้องานวิจัย การศึกษาการใช้ปัญญาประดิษฐ์เพื่อพัฒนาเมทาดาทาภาพถ่ายดิจิทัล
ชื่อผู้วิจัย ผู้ช่วยศาสตราจารย์ ดร. อัครา ธรรมมาสถิตย์กุล
สถาบัน มหาวิทยาลัยบูรพา
คำสำคัญ เมทาดาทา/ภาพถ่าย/เครื่องมือ/ปัญญาประดิษฐ์/ดิจิทัล

บทคัดย่อ

เมทาดาทาภาพถ่าย คือคำที่ใช้อธิบายลักษณะสำคัญต่าง ๆ ที่อยู่ในภาพถ่าย มีส่วนสำคัญในการจัดเก็บคลังข้อมูลภาพถ่าย ทั้งช่วยในการทำดัชนี การจัดหมวดหมู่ กำหนดคำค้น รวมถึงเพิ่มประสิทธิภาพการค้นหา ด้วยความต้องการที่เพิ่มขึ้นสำหรับการจัดการภาพถ่าย และกำหนดเมทาดาทาแบบอัตโนมัติได้มีโปรแกรมที่สามารถใช้ในการกำหนดเมทาดาทาภาพถ่ายอัตโนมัติได้หลายโปรแกรม งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษา และประเมินประสิทธิภาพเครื่องมือปัญญาประดิษฐ์ที่สามารถรู้จำวัตถุในภาพถ่ายได้ 4 โปรแกรมได้แก่ อเมซอน-เรคคอคนิชัน คลาริฟาย อิมเมกกา และกูเกิลคลาวน์วิชันเอพีไอ เพื่อวิเคราะห์โปรแกรมที่เหมาะสมสำหรับการกำหนดเมทาดาทาสำหรับภาพถ่ายแบบอัตโนมัติ

การศึกษาได้เปรียบเทียบประสิทธิภาพของเครื่องมือเหล่านี้บนชุดข้อมูลของรูปภาพซึ่งครอบคลุมหมวดหมู่ต่าง ๆ ได้แก่ คน สัตว์ ต้นไม้-ดอกไม้ วิวทิวทัศน์ ผัก-ผลไม้ อาหาร ยานพาหนะ สถานที่ท่องเที่ยว ศิลปวัฒนธรรม และภาพปกหนังสือ/โปสเตอร์เก่า ด้วยการวัดค่าคำทับซ้อน การวัดค่าคำที่ไม่ทับซ้อน การวัดค่าความคล้ายโคไซน์ การหาค่าความไว ค่าความแม่นยำ และ ค่าความถูกต้องเอฟวัน

ผลลัพธ์แสดงให้เห็นว่าเครื่องมือทั้งหมดสามารถใช้ในการกำหนดเมทาดาทาได้ในประสิทธิภาพที่แตกต่างกัน จากการศึกษาพบว่าเครื่องมือที่ดีที่สุดสำหรับการกำหนดรูปภาพของคน สัตว์ และอาหารคือ โปรแกรมคลาริฟาย โปรแกรมอเมซอน เรคคอคนิชัน ทำงานได้ดีที่สุดกับภาพผักผลไม้ และยานพาหนะ โปรแกรมอิมเมกกา ทำงานได้ดีที่สุดกับภาพต้นไม้ดอกไม้ ศิลปวัฒนธรรม และปกหนังสือ/โปสเตอร์เก่า ในขณะที่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ ทำงานได้ดีที่สุดกับภาพถ่ายวิวทิวทัศน์ และสถานที่ท่องเที่ยว ถึงแม้ว่าเครื่องมือปัญญาประดิษฐ์นี้จะสามารถกำหนดเมทาดาทาได้ แต่การตรวจสอบจากมนุษย์ในขั้นตอนสุดท้ายยังเป็นสิ่งที่จำเป็นเพื่อให้ได้เมทาดาทาที่สมบูรณ์ และถูกต้องมากที่สุด

Title A study of using artificial intelligence to enhance digital image metadata
Author Asst. Dr. Akara Thammastitkul
Institution Burapha University
Keyword metadata/image/tool/artificial intelligence/digital

Abstract

Image metadata known as text data contains information that is relevant to or can identify objects in the image. Effective metadata aid in managing images, including indexing, categorizing, and creating search terms. With the increasing demand for image manipulation and automatically generate metadata has led to the development of numerous image recognition tools. In this research report, we evaluate four image recognition tools: Amazon Rekognition, Clarifai, Imagga, and Google Cloud Vision API. Our study aims to analyze the appropriate program for automatic image metadata determination.

We conducted experiments on several datasets consisting of various image categories, such as human, animal, plant and flower, view and landscape, vegetable and fruit, food, vehicle, tourist landmark, art and culture, and old book cover and posters. We used word overlap similarity measure, no word overlap similarity measure, cosine similarity, precision, recall, and F1-score as evaluation metrics to assess the performance of each tool.

The results showed that all of the tools performed well, but there were notable differences in performance across categories. Overall, the best tool for human animal and food image tagging was found to be Clarifai, while Amazon Rekognition performed best for vegetable-fruit and vehicle images. Imagga was found to be the best tool for plant and flower, art and culture, and old book cover and posters image recognition, while Google Cloud Vision API performed best for view and landscape and tourist landmark recognition. These findings highlight the importance of carefully selecting the appropriate tool for a given task and suggest that the performance of these tools may be improved through further training and refinement. Although this Artificial intelligence system is capable of identifying metadata. However, the most accurate and reliable technique to acquire precise metadata is to conduct a final human supervision.

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
สารบัญ ..... ฉ
สารบัญตาราง ..... ซ
สารบัญภาพ ..... ฎ
บทที่ 1 บทนำ 1
ความเป็นมาของปัญหา 1
วัตถุประสงค์ของการวิจัย 2
ขอบเขตของการวิจัย 3
กรอบแนวคิดการวิจัย 4
ประโยชน์ที่คาดว่าจะได้รับ 4
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 5
เมทาดาทาภาพถ่าย (Image Metadata) 5
การจดจำภาพ (Image Recognition) 7
ระบบเรียนรู้เชิงลึก (Deep learning: DL) 10
กระบวนการของระบบจดจำภาพ (Image Recognition Process) 11
ซอฟต์แวร์จดจำรูปภาพ (Image Recognition Software) 12
เอพีไอ (Application Program Interface: API) 14
งานวิจัยที่เกี่ยวข้อง 15
บทที่ 3 วิธีทำการทดลอง 17
เก็บรวบรวมภาพถ่าย 18
บันทึกเมทาดาทาที่กำกับมาพร้อมภาพถ่าย 21
ทดสอบเครื่องมือการรู้จำภาพถ่ายด้วยเทคนิคปัญญาประดิษฐ์ 23
บันทึกเมทาดาทาที่ได้จากการใช้เครื่องมือ 50
การวิเคราะห์ประสิทธิภาพ 54

สารบัญ (ต่อ)

หน้า

บทที่ 4 ผลการทดลอง 59
ผลการเก็บรวบรวมภาพถ่าย พร้อมบันทึกเมทาดาทาที่มาพร้อมภาพถ่าย 59
ผลการทดสอบเครื่องมือปัญญาประดิษฐ์ และบันทึกเมทาดาทาที่ได้จากเครื่องมือ 60
ผลการคำนวณค่าความถูกต้องเมทาดาทา และประสิทธิภาพเครื่องมือ 90
ผลการทดลองเครื่องมือปัญญาประดิษฐ์ที่เหมาะสมกับภาพถ่ายแต่ละประเภท 103
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ 107
สรุป 107
อภิปรายผล 109
ข้อเสนอแนะ 121
เอกสารอ้างอิง 122

สารบัญตาราง

หน้า

ตารางที่ 1 เมทาดาทาที่ติดมากับข้อมูลหลัก 5
ตารางที่ 2 โครงสร้างข้อมูล IPTC 8
ตารางที่ 3 แบบบันทึกข้อมูลเมทาดาทาที่กำกับมาพร้อมภาพถ่าย 22
ตารางที่ 4 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมอเมซอน เรคคอคนิชันของภาพรหัส A010 50
ตารางที่ 5 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมอิมเมกกาของภาพรหัส A010 50
ตารางที่ 6 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมคลาริฟายของภาพรหัส A010 51
ตารางที่ 7 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมกูเกิลคลาวน์วิชันเอพีไอของภาพรหัส A010 52
ตารางที่ 8 เปรียบเทียบเมทาดาทาเรียงลำดับตามค่าความเชื่อมั่น 53
ตารางที่ 9 เปรียบเทียบเมทาดาทาเรียงลำดับตามตัวอักษร 53
ตารางที่ 10 ผลการวัดค่าคำที่ทับซ้อนของภาพรหัส A010 54
ตารางที่ 11 ผลการวัดค่าคำที่ไม่ทับซ้อนของภาพรหัส A010 55
ตารางที่ 12 ผลการวัดค่าความคล้ายโคไซน์ ของภาพรหัส A010 56
ตารางที่ 13 ผลการวัดค่าความถูกต้องเมทาดาทาของภาพรหัส A010 58
ตารางที่ 14 เมทาดาทาจากการทดลองในหมวดหมู่ภาพคน รหัสภาพ H01 60
ตารางที่ 15 เมทาดาทาจากการทดลองในหมวดหมู่ภาพคน รหัสภาพ H02 61
ตารางที่ 16 เมทาดาทาจากการทดลองในหมวดหมู่ภาพคน รหัสภาพ H11 62
ตารางที่ 17 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสัตว์ รหัสภาพ A04 63
ตารางที่ 18 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสัตว์ รหัสภาพ A06 64
ตารางที่ 19 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสัตว์ รหัสภาพ A11 65
ตารางที่ 20 เมทาดาทาจากการทดลองในหมวดหมู่ภาพต้นไม้ ดอกไม้ รหัสภาพ PF01 66
ตารางที่ 21 เมทาดาทาจากการทดลองในหมวดหมู่ภาพต้นไม้ ดอกไม้ รหัสภาพ PF02 67
ตารางที่ 22 เมทาดาทาจากการทดลองในหมวดหมู่ภาพต้นไม้ ดอกไม้ รหัสภาพ PF04 68
ตารางที่ 23 เมทาดาทาจากการทดลองในหมวดหมู่ภาพวิวทิวทัศน์ รหัสภาพ VL05 69
ตารางที่ 24 เมทาดาทาจากการทดลองในหมวดหมู่ภาพวิวทิวทัศน์ รหัสภาพ VL07 70
ตารางที่ 25 เมทาดาทาจากการทดลองในหมวดหมู่ภาพวิวทิวทัศน์ รหัสภาพ VL09 71
ตารางที่ 26 เมทาดาทาจากการทดลองในหมวดหมู่ภาพผัก ผลไม้ รหัสภาพ VF04 72
ตารางที่ 27 เมทาดาทาจากการทดลองในหมวดหมู่ภาพผัก ผลไม้ รหัสภาพ VF05 73
ตารางที่ 28 เมทาดาทาจากการทดลองในหมวดหมู่ภาพผัก ผลไม้ รหัสภาพ VF06 74
ตารางที่ 29 เมทาดาทาจากการทดลองในหมวดหมู่ภาพอาหาร รหัสภาพ F07 75
ตารางที่ 30 เมทาดาทาจากการทดลองในหมวดหมู่ภาพอาหาร รหัสภาพ F09 76

สารบัญตาราง (ต่อ)

หน้า

ตารางที่ 31 เมทาดาทาจากการทดลองในหมวดหมู่ภาพอาหาร รหัสภาพ F15 77
ตารางที่ 32 เมทาดาทาจากการทดลองในหมวดหมู่ภาพยานพาหนะ รหัสภาพ V03 78
ตารางที่ 33 เมทาดาทาจากการทดลองในหมวดหมู่ภาพยานพาหนะ รหัสภาพ V07 79
ตารางที่ 34 เมทาดาทาจากการทดลองในหมวดหมู่ภาพยานพาหนะ รหัสภาพ V10 80
ตารางที่ 35 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสถานที่ท่องเที่ยว รหัสภาพ L01 81
ตารางที่ 36 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสถานที่ท่องเที่ยว รหัสภาพ L06 82
ตารางที่ 37 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสถานที่ท่องเที่ยว รหัสภาพ L15 83
ตารางที่ 38 เมทาดาทาจากการทดลองในหมวดหมู่ภาพศิลปวัฒนธรรม รหัสภาพ AC05 84
ตารางที่ 39 เมทาดาทาจากการทดลองในหมวดหมู่ภาพศิลปวัฒนธรรม รหัสภาพ AC10 85
ตารางที่ 40 เมทาดาทาจากการทดลองในหมวดหมู่ภาพศิลปวัฒนธรรม รหัสภาพ AC30 86
ตารางที่ 41 เมทาดาทาจากการทดลองในหมวดหมู่ภาพปกหนังสือ/โปสเตอร์เก่า รหัสภาพ BP02 87
ตารางที่ 42 เมทาดาทาจากการทดลองในหมวดหมู่ภาพปกหนังสือ/โปสเตอร์เก่า รหัสภาพ BP03 88
ตารางที่ 43 เมทาดาทาจากการทดลองในหมวดหมู่ภาพปกหนังสือ/โปสเตอร์เก่า รหัสภาพ BP07 89
ตารางที่ 44 เมทาดาทาเรียงตามลำดับตัวอักษรจากตัวอย่างภาพคน 90
ตารางที่ 45 เมทาดาทาทับซ้อนของตัวอย่างภาพคน 91
ตารางที่ 46 เมทาดาทาที่มีความหมายเกี่ยวข้องของตัวอย่างภาพคน 91
ตารางที่ 47 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพคน 92
ตารางที่ 48 ค่าประสิทธิภาพของเมทาดาทาภาพในหมวดภาพคน 92
ตารางที่ 49 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพสัตว์ 93
ตารางที่ 50 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพสัตว์ 93
ตารางที่ 51 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพต้นไม้ ดอกไม้ 94
ตารางที่ 52 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพต้นไม้ ดอกไม้ 94
ตารางที่ 53 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพวิวทิวทัศน์ 95
ตารางที่ 54 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพวิวทิวทัศน์ 95
ตารางที่ 55 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพผัก ผลไม้ 96
ตารางที่ 56 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพผัก ผลไม้ 96
ตารางที่ 57 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพอาหาร 97
ตารางที่ 58 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพอาหาร 97
ตารางที่ 59 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพยานพาหนะ 98
ตารางที่ 60 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพยานพาหนะ 98

สารบัญตาราง (ต่อ)

หน้า

ตารางที่ 61 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพสถานที่ท่องเที่ยว 99
ตารางที่ 62 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพสถานที่ท่องเที่ยว 99
ตารางที่ 63 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพศิลปวัฒนธรรม 100
ตารางที่ 64 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพศิลปวัฒนธรรม 100
ตารางที่ 65 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพปกหนังสือ/โปสเตอร์เก่า..... 101
ตารางที่ 66 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพปกหนังสือ/โปสเตอร์เก่า..... 101
ตารางที่ 67 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาจากการทดลองทั้งหมด 1,000 ภาพ..... 102
ตารางที่ 68 ค่าประสิทธิภาพเฉลี่ยจากการทดลองภาพทั้งหมด 1,000 ภาพ 102
ตารางที่ 69 สรุปเครื่องมือที่เหมาะสมกับภาพถ่ายแต่ละประเภท 105
ตารางที่ 70 ลำดับเครื่องมือที่เหมาะสมกับภาพถ่ายแต่ละประเภท..... 106
ตารางที่ 71 โปรแกรมที่ให้ค่าความถูกต้องสูงสุดของแต่ละหมวดภาพ 106

สารบัญภาพ

หน้า

ภาพที่ 1 กรอบแนวความคิดการวิจัย 4
ภาพที่ 2 กระบวนการเพิ่มข้อมูลเมทาดาทาในรูปถ่าย 6
ภาพที่ 3 การใส่เมทาดาทาด้วยโปรแกรม XnViewMP 7
ภาพที่ 4 การรู้จำวัตถุในภาพถ่าย (ที่มา: https://aws.amazon.com/th/rekognition/) 9
ภาพที่ 5 ความสัมพันธ์ระหว่างปัญญาประดิษฐ์ แมชชีนเลิร์นนิง และ การเรียนรู้เชิงลึก 10
ภาพที่ 6 ขั้นตอนการทำงานของการเรียนรู้เชิงลึก 10
ภาพที่ 7 ขั้นตอนการจดจำภาพ 11
ภาพที่ 8 การทำงานของเอพีไอ 14
ภาพที่ 9 ขั้นตอนการวิจัย 17
ภาพที่ 10 ตัวอย่างภาพถ่ายหมวดหมู่คน 19
ภาพที่ 11 ตัวอย่างภาพถ่ายหมวดหมู่สัตว์ 19
ภาพที่ 12 ตัวอย่างภาพถ่ายหมวดหมู่ต้นไม้ ดอกไม้ 19
ภาพที่ 13 ตัวอย่างภาพถ่ายหมวดหมู่วิวทิวทัศน์ 19
ภาพที่ 14 ตัวอย่างภาพถ่ายหมวดหมู่ผัก ผลไม้ 20
ภาพที่ 15 ตัวอย่างภาพถ่ายหมวดหมู่อาหาร 20
ภาพที่ 16 ตัวอย่างภาพถ่ายหมวดหมู่ยานพาหนะ 20
ภาพที่ 17 ตัวอย่างภาพถ่ายหมวดหมู่สถานที่ท่องเที่ยว 20
ภาพที่ 18 ตัวอย่างภาพถ่ายหมวดหมู่ศิลปวัฒนธรรม 21
ภาพที่ 19 ตัวอย่างภาพถ่ายหมวดหมู่ปกหนังสือ/โปสเตอร์เก่า 21
ภาพที่ 20 ตัวอย่างภาพถ่ายรหัส H001 PF001 และ A010 21
ภาพที่ 21 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมอเมซอน เรคคอคนิชัน 25
ภาพที่ 22 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมคลาริฟาย 31
ภาพที่ 23 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมอิมเมกกา 35
ภาพที่ 24 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมอิมเมกกาแบบยาวเป็นภาษาอังกฤษ 36
ภาพที่ 25 ตัวอย่างผลการแยกหมวดหมู่ภาพถ่ายด้วยโปรแกรมอิมเมกกาเป็นภาษาไทย 37
ภาพที่ 26 ตัวอย่างการแยกหมวดหมู่ด้วยโปรแกรมอิมเมกกา 37
ภาพที่ 27 ตัวอย่างผลการตรวจจับวัตถุในภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ 47
ภาพที่ 28 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ 47
ภาพที่ 29 ตัวอย่างผลการรู้จำสีในภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ 48
ภาพที่ 30 ตัวอย่างผลการรู้จำความปลอดภัยในภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ 48

สารบัญภาพ (ต่อ)

หน้า

ภาพที่ 31 การเชื่อมโยงผลลัพธ์จากโปรแกรมกูเกิลคลาวน์วิชันเอพีไอไปยังโปรแกรมกูเกิล 49
ภาพที่ 32 ผลค่าคำทับซ้อนของภาพแต่ละประเภท 103
ภาพที่ 33 ผลค่าคำที่ไม่ทับซ้อนของภาพแต่ละประเภท 103
ภาพที่ 34 ผลการวัดค่าความคล้ายโคไซน์ของภาพแต่ละประเภท 104
ภาพที่ 35 ผลค่าความถูกต้องเอฟวันของภาพแต่ละประเภท 104
ภาพที่ 36 เปรียบเทียบเมทาดาทาภาพผลไม้จากโปรแกรม (ก) อเมซอน เรคคอคนิชัน และ (ข) อิมเมกกา 111
ภาพที่ 37 เปรียบเทียบเมทาดาทาภาพยานพาหนะจากโปรแกรม (ก) อเมซอน เรคคอคนิชัน และ (ข) กูเกิลคลาวน์วิชันเอพีไอ 111
ภาพที่ 38 เปรียบเทียบเมทาดาทาภาพคนจากโปรแกรม (ก) คลาริฟาย และ (ข) กูเกิลคลาวน์วิชันเอพีไอ 112
ภาพที่ 39 เปรียบเทียบเมทาดาทาภาพสัตว์จากโปรแกรม (ก) คลาริฟาย และ (ข) กูเกิลคลาวน์วิชันเอพีไอ 112
ภาพที่ 40 เปรียบเทียบเมทาดาทาภาพอาหารจากโปรแกรม (ก) คลาริฟาย และ (ข) กูเกิลคลาวน์วิชันเอพีไอ 113
ภาพที่ 41 เปรียบเทียบเมทาดาทาภาพดอกไม้จากโปรแกรม (ก) อิมเมกกา และ (ข) กูเกิลคลาวน์วิชันเอพีไอ 114
ภาพที่ 42 ผลการกำหนดเมทาดาทาภาษาไทยของภาพดอกไม้จากโปรแกรมอิมเมกกา 114
ภาพที่ 43 ผลการกำหนดเมทาดาทาภาพศิลปวัฒนธรรม ทั้งภาษาอังกฤษ และภาษาไทยจากโปรแกรมอิมเมกกา 115
ภาพที่ 44 ผลเมทาดาทา และตำแหน่งที่ตั้ง ภาพศิลปวัฒนธรรมจากโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ 115
ภาพที่ 45 เปรียบเทียบเมทาดาทาภาพโปสเตอร์จากโปรแกรม (ก)อิมเมกกา และ (ข)กูเกิลคลาวน์วิชันเอพีไอ 116
ภาพที่ 46 เปรียบเทียบเมทาดาทาภาพวิวทิวทัศน์จากโปรแกรม (ก) กูเกิลคลาวน์วิชันเอพีไอ และ (ข) อเมซอน เรคคอคนิชัน 117
ภาพที่ 47 ผลการใช้โปรแกรมกูเกิลคลาวน์วิชันเอพีไอกับภาพสถานที่ท่องเที่ยวลอนดอนอาย (The London eye) (ก) ตำแหน่งสถานที่ (ข) วัตถุ (ค) เมทาดาทา (ฆ) ตัวอักษร (ง) คุณสมบัติ และ (จ) ความเหมาะสม 118
ภาพที่ 48 ผลลัพธ์การใช้โปรแกรมกูเกิลคลาวน์วิชันในการแยกข้อความจากภาพ 119

บทที่ 1
บทนำ

ความเป็นมาของปัญหา

ในยุคดิจิทัลที่มีความก้าวหน้าอย่างรวดเร็วของเทคโนโลยี ทำให้เกิดสื่อดิจิทัลขึ้นจำนวนมาก ทั้งอยู่ในรูปแบบของข้อความ ภาพนิ่ง เสียง และวีดีโอ โดยเฉพาะจำนวนผู้ใช้โทรศัพท์มือถือที่เพิ่มขึ้นอย่างมากในปี พ.ศ. 2564 ผลสำรวจพบประชากรไทยที่อายุ 6 ปีขึ้นไป มีผู้ใช้โทรศัพท์มือถือถึง 60.6 ล้านคน จากจำนวนประชากรทั้งหมดประมาณ 64 ล้านคน นับเป็น 94.70% โดยกิจกรรมที่เกิดมากที่สุดคือการใช้โซเชียลมีเดียโพสกิจกรรมผ่านทางภาพถ่าย ถึง 78.70% (สำนักงานสถิติแห่งชาติ, 2564) การเกิดสื่อประเภทรูปภาพ หรือภาพถ่ายจำนวนมากนี้ หากไม่มีการจัดการที่ดี จะทำให้ยากต่อการเข้าถึง และใช้งานในอนาคต การจัดการข้อมูลสารสนเทศภาพถ่ายดิจิทัลจึงเป็นกระบวนการสำคัญ

ในการพัฒนาคลังข้อมูลดิจิทัลของทุกหน่วยงานการจัดการข้อมูลสื่อสารสนเทศดิจิทัล ต้องอาศัยเครื่องมือ หรือคำอธิบายรายการที่เรียกว่า เมทาดาทา (metadata) ในการลงรายการข้อมูลที่บ่งชี้คุณลักษณะเฉพาะของสารสนเทศนั้น ๆ โดยที่สื่อสารสนเทศแต่ละประเภทจะมีมาตราฐานการลงเมทาดาทาที่แตกต่างกัน (Witten et al., 2010) ในงานวิจัยนี้สนใจศึกษาสารสนเทศประเภทภาพถ่าย ซึ่งเมทาดาทาสำหรับภาพถ่ายจะแบ่งเป็น 2 ส่วน ประกอบด้วย ส่วนที่ 1 ข้อมูล EXIF (Exchangeable Image File Format) และ ส่วนที่ 2 ข้อมูล IPTC/XMP (International Press Telecommunications Council/ Extensible Metadata Platform) (International Press Telecommunications Council, 2022) ส่วนที่ 1 ข้อมูล EXIF เป็นเมทาดาทาที่ให้ข้อมูลเฉพาะเกี่ยวกับรูปถ่าย พัฒนาขึ้นโดยสมาคมอุตสาหกรรมอิเล็กทรอนิกส์และเทคโนโลยีสารสนเทศแห่งญี่ปุ่น (The Japan Electronic Industry Development Association: JEIDA) ซึ่งข้อมูลต่าง ๆ จะถูกบรรจุลงในภาพที่ถ่ายโดยกล้องดิจิทัล เป็นข้อมูลที่ติดมากับข้อมูลหลัก เช่น รหัสภาพ การตั้งค่าของกล้อง ขนาดภาพ ความละเอียด ประเภทไฟล์ภาพ และ ส่วนที่ 2 ข้อมูล IPTC/XMP เป็นมาตรฐานการจัดเก็บข้อมูลลงในไฟล์ภาพ คือข้อมูลทางกายภาพในภาพถ่ายที่ผู้ใช้กำหนดเอง เช่นชื่อภาพ สถานที่ เจ้าของภาพ สถานที่ คำอธิบายภาพ การสร้างข้อมูลเมทาดาทาที่กำหนดเอง เป็นการระบุสิ่งที่มีอยู่ในภาพด้วยคำค้น หรือคีย์เวิร์ดที่คลอบคลุมสำหรับการค้นหาในอนาคต การกำหนดเมทาดาทาให้ครอบคลุมสำหรับการค้นหาจึงเป็นสิ่งที่จำเป็นต้องทำ การกำหนดเมทาดาทาด้วยแรงงานมนุษย์นี้ต้องใช้เวลา และทักษะความชำนาญของผู้ให้ข้อมูล การบันทึกเมทาดาทาที่ดีจะนำไปสู่การจัดหมวดหมู่และค้นหาข้อมูลที่ดีมากยิ่งขึ้น (Mustafa et al., 2021) ด้วยปัญหาของปริมาณภาพถ่ายที่เกิดขึ้นจำนวนมาก การกำหนดเมทาดาทาแบบอัตโนมัติจึงได้รับความนิยมมากขึ้น การใช้เทคโนโลยีเพื่อกำหนดเมทาดาทาแบบอัตโนมัตินี้นอกจากจะช่วยให้ประหยัดเวลาแล้ว ยังก่อให้เกิดผลงานที่มีประสิทธิภาพมากยิ่งขึ้นด้วย (Bunn, 2021; Jailiant & Caputo, 2022)

เทคโนโลยีปัญญาประดิษฐ์ (Artificial intelligent: AI) เป็นเครื่องมือทางคอมพิวเตอร์ที่มีความสามารถในการเรียนรู้ และเข้าใจองค์ความรู้ต่าง ๆ โดยใช้ข้อมูลที่มีอยู่มาประกอบกับการประมวลผลขั้นสูง นำมาใช้ในหน่วยงานหลายประเภทอย่างแพร่หลาย เช่น หุ่นยนต์ การทำนายพยากรณ์ การรู้จำ การแปลภาษา การโต้ตอบกับผู้ช่วยเสมือนจริง ระบบช่วยเหลือ และการแนะนำสินค้า ซึ่งการประมวลผลภาพ หรือการวิเคราะห์ภาพ (Image processing) นั้นถือเป็นหนึ่งในเทคโนโลยีที่สำคัญเกี่ยวกับปัญญาประดิษฐ์ ในการรู้จำวัตถุในภาพถ่าย ที่สามารถนำมาใช้ในการกำหนดเมทาดาทาอัตโนมัติจากภาพถ่ายดิจิทัล ลดเวลาในการทำงานและทำให้ได้เมทาดาทาที่สมบูรณ์มากยิ่งขึ้น ในปัจจุบันมีการพัฒนาเครื่องมือการรู้จำวัตถุในภาพถ่ายด้วยปัญญาประดิษฐ์ขึ้นในหลายแพลตฟอร์ม ทั้งซอฟต์แวร์แบบเปิด (open source) ฟรีซอฟต์แวร์ (freeware) และแบบลิขสิทธิ์ เสียค่าใช้จ่าย (license) เช่น แบรนด์วอช (Brandwatch) อิมเมจ อินไซด์ (Image Insights) โลโคบัซ (Locobuzz) ไมโครซอฟท์ แอซเซอะ (Microsoft azure) อเมซอน เรคคอคนิชัน (Amazon Rekognition) คลาริฟาย (Clarifai) อิมเมกกา (Imagga) และกูเกิลคลาวน์วิชันเอพีไอ (Google cloud vision API) ซึ่งแต่ละแพลตฟอร์มมีจุดเด่นที่เหมาะกับการใช้งานเฉพาะตัวที่แตกต่างกัน การศึกษาและค้นหาแพลตฟอร์มที่เหมาะสมสำหรับการกำหนดเมทาดาทาสำหรับภาพถ่าย จะสามารถช่วยในการตัดสินใจเลือกแพลตฟอร์มที่เหมาะสมสำหรับการกำหนดเมทาดาทาแบบอัตโนมัติให้ครอบคลุมสาระตามลักษณะเฉพาะของภาพถ่ายที่หน่วยงานจัดเก็บ เพิ่มศักยภาพในการทำคลังข้อมูลภาพดิจิทัล เพิ่มประสิทธิภาพการทำระบบจัดเก็บ การเรียกใช้และค้นคืนรูปภาพ รวมถึงแลกเปลี่ยนข้อมูลกับระบบคลังข้อมูลอื่น ๆ ได้ต่อไปในอนาคต

จากการทบทวนวรรณกรรมพบว่าปัจจุบันยังคงมีงานวิจัยทางด้านการใช้ปัญญาประดิษฐ์มาช่วยพัฒนาเมทาดาทาจากภาพถ่ายอยู่น้อย ไม่มีการวัดประสิทธิภาพของเมทาดาทาที่ได้ และประสิทธิภาพของเครื่องมือที่ชัดเจน ทำให้ไม่สามารถให้คำแนะนำการใช้เครื่องมือที่เหมาะสมกับรูปภาพที่มีลักษณะแตกต่างกันได้ งานวิจัยนี้จะช่วยทำการวิเคราะห์ และทดสอบเครื่องมือในการกำหนดเมทาดาทาแบบอัตโนมัติ ซึ่งจะช่วยให้การกำหนดคำค้นมีประสิทธิภาพและรวดเร็วมากยิ่งขึ้น เอื้อต่อการเก็บข้อมูลอย่างเป็นระบบ และสามารถเข้าถึงได้ง่ายในอนาคต สามารถนำองค์ความรู้ที่ได้ประยุกต์ใช้ในการเก็บข้อมูลภาพถ่ายแบบส่วนตัว องค์กร รวมถึงสถาบันคลังข้อมูลดิจิทัล หรือห้องสมุดดิจิทัลได้

วัตถุประสงค์ของการวิจัย

1. เพื่อสำรวจเครื่องมือในการรู้จำวัตถุในภาพถ่ายที่ใช้ปัญญาประดิษฐ์เพื่อกำหนดเมทาดาทา
2. เพื่อวิเคราะห์และประเมินประสิทธิภาพเครื่องมือในการรู้จำภาพที่เหมาะสมกับภาพถ่ายแต่ละประเภท

ขอบเขตของการวิจัย

ขอบเขตด้านเนื้อหา

งานวิจัยนี้เป็นการศึกษาการใช้ปัญญาประดิษฐ์เพื่อพัฒนาเมทาดาทาภาพถ่ายดิจิทัล โดยทำการรวบรวมภาพถ่ายดิจิทัล ทดสอบการรู้จำวัตถุในภาพถ่าย วัดค่าความถูกต้อง แล้วนำมาวิเคราะห์และประเมินประสิทธิภาพผลเมทาดาทาที่ได้จากเครื่องมือในการรู้จำเปรียบเทียบกับเมทาดาทาที่กำกับมาพร้อมภาพถ่าย

กลุ่มตัวอย่างที่ใช้ศึกษา

1. ภาพถ่ายที่มีเมทาดาทากำกับ จากคลังภาพดิจิทัลแบบเปิด เผยแพร่ภายใต้โดเมนสาธารณะ (public domain) ซึ่งไม่มีลิขสิทธิ์ หรือเจ้าของลิขสิทธิ์ยกให้เป็นสาธารณะประโยชน์ จำนวน 1,000 ภาพ จากเว็บไซต์ครีเอทีฟคอมมอนส์ (https://creativecommons.org/) และฟลิคเกอร์ (https://www.flickr.com/creativecommons/)

คุณสมบัติของภาพถ่ายที่นำมาใช้ประกอบด้วย

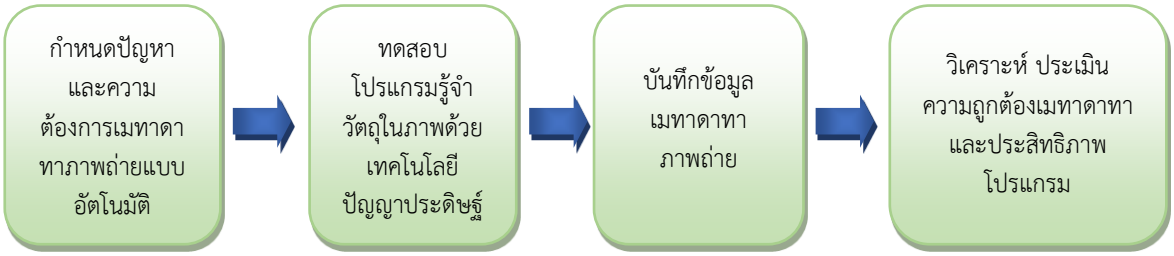
1. ภาพสี ประเภทไฟล์ .jpg หรือ .png
2. ความละเอียดไม่น้อยกว่า 80 พิกเซล x 80 พิกเซล และไม่มากกว่า 4096 พิกเซล x 4096 พิกเซล
3. ขนาดไฟล์ไม่มากกว่า 4 เมกะไบต์
4. มีเมทาดาทากำกับไม่น้อยกว่า 5 คำ และไม่มากกว่า 50 คำ
5. ภาพประกอบด้วยคน สัตว์ สิ่งของ อาหาร ธรรมชาติ หรือสถานที่

2. เครื่องมือการรู้จำภาพถ่ายด้วยเทคนิคปัญญาประดิษฐ์จำนวน 4 โปรแกรม คือ

1. อเมซอน เรคคอคนิชัน (Amazon Rekognition)
2. คลาริฟาย (Clarifai)
3. อิมเมกกา (Imagga) และ
4. กูเกิลคลาวน์วิชันเอพีไอ (Google cloud vision API)

โดยมีเกณฑ์การพิจารณาเลือกเครื่องมือจากความสามารถโดดเด่นในการรู้จำวัตถุในภาพ ทำงานบนระบบคลาวน์ มีตัวกลางการเรียกใช้โปรแกรมหรือเอพีไอ ติดตั้งใช้งานง่าย และมีการพัฒนาโปรแกรมอย่างต่อเนื่อง

กรอบแนวคิดการวิจัย



ภาพที่ 1 กรอบแนวความคิดการวิจัย

ประโยชน์ที่คาดว่าจะได้รับ

งานวิจัยนี้จะเป็นประโยชน์ต่อการจัดการภาพถ่ายดิจิทัล การทำคลังข้อมูลดิจิทัล ที่การกำหนดเมทาดาทาที่สมบูรณ์เป็นขั้นตอนที่สำคัญและมีความจำเป็น กระบวนการกำหนดเมทาดาทาแบบอัตโนมัติจากเครื่องมือทางปัญญาประดิษฐ์จะช่วยให้ได้เมทาดาทาที่ครอบคลุมเนื้อหาสาระในภาพถ่าย ด้วยคำค้นที่หลากหลายมากขึ้น เป็นแนวทางในการตัดสินใจใช้เครื่องมือในการกำหนดเมทาดาทาภาพถ่ายให้เหมาะสมกับลักษณะของภาพถ่าย ช่วยให้การทำงานของผู้จัดการข้อมูลภาพถ่ายสะดวกและมีประสิทธิภาพมากยิ่งขึ้น

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เมทาดาทาภาพถ่าย (Image Metadata)

เมทาดาทา (Metadata) หมายถึง ข้อมูลที่ใช้บ่งชี้บอกคุณลักษณะ หรือให้ข้อมูลเกี่ยวกับข้อมูล ซึ่งมีอยู่มากมายหลายลักษณะ สำหรับงานห้องสมุด ได้นำเมทาดาทามาช่วยในการลงรายการข้อมูลของทรัพยากรสารสนเทศที่เพิ่มจำนวนขึ้นอย่างรวดเร็ว ทั้งทรัพยากรสารสนเทศที่เป็นดิจิทัล และไม่เป็นดิจิทัล เนื่องจากเมทาดาทาช่วยเพิ่มความสามารถในการค้นคืน และเข้าถึงทรัพยากรสารสนเทศ ช่วยสงวนรักษาข้อมูล ขยายขอบเขตการใช้งานสารสนเทศ และบ่งชี้คุณสมบัติเฉพาะของสารสนเทศ มาตรฐานการลงรายการต่าง ๆ ที่นำมาใช้ในงานห้องสมุด จึงทำให้เกิดความเข้าใจที่ตรงกัน ทำให้สะดวก และไม่เกิดความสับสนในการใช้ทรัพยากรสารสนเทศ (Witten et al., 2010) เมทาดาทาภาพถ่ายจะแบ่งเป็น 2 ส่วน คือ ข้อมูลที่ติดมากับข้อมูลหลัก เช่น รหัสภาพ การตั้งค่าของกล้อง ขนาดภาพ ความละเอียด ประเภทไฟล์ภาพ และข้อมูลทางกายภาพในภาพถ่ายที่ผู้ใช้กำหนดเอง เช่น ชื่อภาพ สถานที่ เจ้าของภาพ สถานที่ คำอธิบายภาพ ข้อมูลรูปแบบไฟล์ภาพที่แลกเปลี่ยนได้ (Exchangeable Image File Format: EXIF) จะถูกบันทึกโดยกล้องโดยอัตโนมัติในเวลาที่ถ่ายหรือสร้างขึ้น ข้อมูลที่เก็บไว้ในรูปถ่ายดิจิทัลอาจรวมถึงวันที่และเวลาที่ถ่ายภาพประเภทและขนาดของไฟล์รูปภาพการตั้งค่ากล้องถ่ายรูปหรือหากใช้กล้องถ่ายรูปหรือโทรศัพท์ที่มีคุณสมบัติจีพีเอส (GPS) ตำแหน่งทางภูมิศาสตร์ (Harran et al., 2017; Jaffe, 2020) ตัวอย่างข้อมูล EXIF แสดงในตารางที่ 1

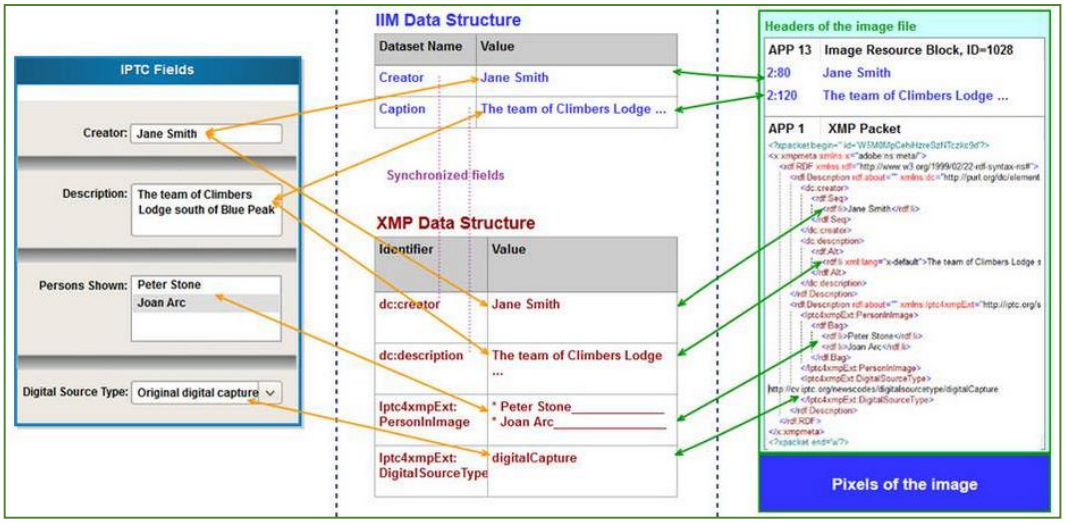
ตารางที่ 1 เมทาดาทาที่ติดมากับข้อมูลหลัก

หัวข้อ	**คำอธิบาย**
File Size	ขนาดไฟล์ข้อมูล
File Type	ประเภทไฟล์ข้อมูล
MIME Type	ป้ายกำกับระบุประเภทข้อมูล
Image Width	ความกว้างภาพถ่าย
Image Height	ความสูงภาพถ่าย
Encoding Process	การเข้ารหัส
Bits Per Sample	ความละเอียดภาพถ่าย หรือจำนวนบิท
Color Components	ส่วนประกอบสี
X Resolution	ความละเอียดภาพแนวนอน
Y Resolution	ความละเอียดภาพแนวตั้ง
YCbCr Sub Sampling	ระบบการเก็บข้อมูลสีแบบดิจิทัล

ที่มา: EXIF Data Viewer (https://exifdata.com)

เมทาดาทาที่ผู้ใช้กำหนด หรือข้อมูล IPTC/XMP (International Press Telecommunications Council/Extensible Metadata Platform) คือข้อมูลที่สามารถแก้ไขได้ อนุญาตให้ผู้ใช้เพิ่มข้อมูลบรรยายในรูปดิจิตอล ช่วยให้สามารถเพิ่มและจัดเก็บข้อมูลรูปภาพ เช่นคำอธิบายภาพ แท็กคำอธิบายข้อมูลลิขสิทธิ์เป็นต้น IPTC คือข้อมูลเฉพาะของรูปภาพทั่วไป เป็นมาตรฐานอุตสาหกรรมที่ใช้กันแพร่หลายมากที่สุดซึ่งเดิมสร้างขึ้นโดย International Press สภาโทรคมนาคมเพื่อเพิ่มข้อมูลเฉพาะภาพรวมทั้งผู้สร้างรายละเอียดและข้อมูลลิขสิทธิ์ได้รับการพัฒนาโดย Adobe ในปี พ.ศ. 2544 จาก IPTC สำหรับวัตถุประสงค์ของผู้ใช้ปลายทางทั้งสองมาตรฐานจะสามารถใช้แทนกันได้

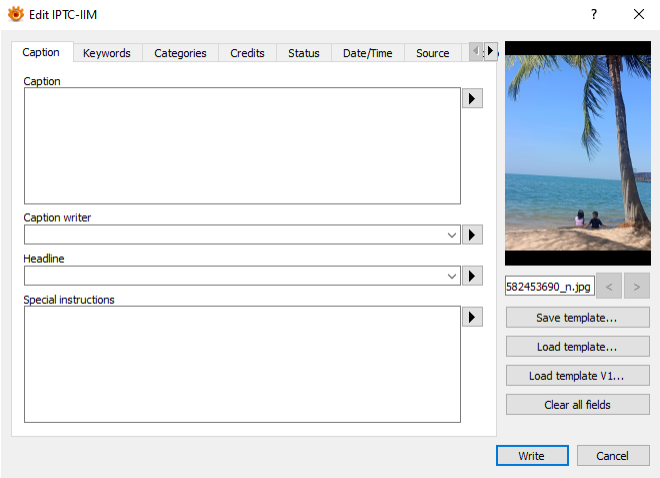
การเพิ่มข้อมูลเมทาดาทาจะช่วยให้ผู้ใช้สามารถใช้ข้อมูลนี้เพื่อจัดระเบียบคอลเลคชันภาพถ่ายได้ โดยสามารถเพิ่มข้อมูลเมทาดาทาได้ด้วยโปรแกรมจัดการรูปภาพทั้งโปรแกรมโอเพนซอร์ส เช่น โปรแกรมอัลเฟรสโก้ (Alfresco) กิมพ (Gimp) เอ็กซิฟทูล (ExifTool) โปรแกรมฟรีแวร์ เช่น เออแฟนวิว (IrfanView) เอ็กเอนวิวเอ็มพี (XnViewMP) และโปรแกรมลิขสิทธิ์ เช่น ไลท์รูม (Lightroom) อะโดบีอิลลัสเตรเตอร์ (Adobe Illustrator) อะโดบี โฟโต้ชอป (Adobe Photoshop) หัวข้อของข้อมูลเมทาดาทาอาจมีความแตกต่างกันขึ้นอยู่กับซอฟต์แวร์ที่ใช้ แต่โดยทั่วไปแล้ว จะประกอบด้วยข้อมูล ผู้เขียน หัวข้อ ลิขสิทธิ์ คำบรรยายภาพ คำค้น ซึ่งเมทาดาทาที่จำเป็นสำหรับการนำไปใช้งานด้านการสืบค้น คือ เจ้าของภาพ (Creator) ชื่อภาพ (Title) คำอธิบาย (Description) และ คำค้น (Keyword) กระบวนการเพิ่มเมทาดาทาแสดงในภาพที่ 2 และตัวอย่างการเพิ่มข้อมูลเมทาดาทาแสดงในภาพที่ 3



แสดงบนคอมพิวเตอร์ ภายในโปรแกรมจัดการ ภายในไฟล์รูปภาพ

ภาพที่ 2 กระบวนการเพิ่มข้อมูลเมทาดาทาในรูปถ่าย

(ปรับปรุงภาพจาก https://iptc.org/std/photometadata/documentation/userguide/)



ภาพที่ 3 การใส่เมทาดาทาด้วยโปรแกรม XnViewMP

โครงสร้างข้อมูล IPTC (IPTC's Structure)

โปรแกรมจัดการภาพถ่าย XnView กำหนดโครงสร้างข้อมูลเมทาดาทาในส่วนของ IPTC ไว้ 8 หมวดหมู่ ประกอบด้วย คำบรรยายภาพ (Caption) คำค้น (Keywords) หมวดหมู่ (Categories) เจ้าของ (Credits) สถานะ (Status) วันที่และเวลา (Date/Time) แหล่งข้อมูล (Source) และ รูปแบบอื่น ๆ (Options) โดยแต่ละหมวดหมู่มีข้อมูลย่อยดังแสดงในตารางที่ 2

การจดจำภาพ (Image Recognition)

เพื่อพัฒนานวัตกรรม และยกระดับทางธุรกิจ ตามแผนยุทธศาสตร์การพัฒนาเชิงพื้นที่ สำนักงานนวัตกรรมแห่งชาติ ได้สนับสนุนให้เกิดการพัฒนาย่านอุตสาหกรรม อารี (ARI INNOVATION DISTRICT) (สำนักงานนวัตกรรมแห่งชาติ องค์การมหาชน, 2564) เพื่อพัฒนาอุตสาหกรรมเชิงลึก ซึ่ง ARI ประกอบไปด้วย A ปัญญาประดิษฐ์ (Artificial intelligence) R เทคโนโลยีหุ่นยนต์ (Robotics) และ I เทคโนโลยีเสมือนจริงและอินเทอร์เน็ตทุกสรรพสิ่ง (Immersive and IOT) มีการประยุกต์ใช้งานปัญญาประดิษฐ์ ในหลายภาคส่วน ทั้งภาคอุตสาหกรรม การค้า การบริการ การศึกษา ในงานวิจัยนี้ได้ประยุกต์ใช้ในภาคสังคม และวัฒนธรรม โดยใช้การวิเคราะห์ภาพ ถือเป็นหนึ่งในเทคโนโลยีฐานที่สำคัญเกี่ยวกับปัญญาประดิษฐ์ ในการรู้จำวัตถุในภาพถ่าย เพื่อช่วยให้การจัดเก็บข้อมูลภาพถ่ายดิจิทัลมีประสิทธิภาพมากขึ้น

ตารางที่ 2 โครงสร้างข้อมูล IPTC

หมวดหมู่	**ข้อมูลย่อย**
คำบรรยายภาพ (Caption)	คำบรรยายภาพ (Caption)
	ผู้ให้คำบรรยาย (Caption writer)
	พาดหัวเรื่อง (Headline)
	คำสั่งพิเศษ (Special instruction)
คำค้น (Keywords)	คำค้น (Keywords)
หมวดหมู่ (Categories)	หมวดหมู่ (Categories)
	หมวดหมู่เพิ่มเติม (Supplemental Categories)
เจ้าของ (Credits)	บรรทัดแรกที่ลงชื่อผู้เขียน (Byline)
	หัวเรื่องบรรทัดแรก(Byline title)
	เจ้าของ (Credits)
	แหล่งข้อมูล (Source)
	ลิขสิทธิ์ (Copyright)
	ข้อมูลติดต่อ (Contact)
สถานะ (Status)	แก้ไขสถานะ (Edit status)
	ความสำคัญ (Priority)
	วงจรวัตถุ (Object cycle)
	หมายเลขงาน (Job ID)
	โปรแกรม (Program)
	เวอร์ชั่นโปรแกรม (Program version)
วันและเวลา (Date/Time)	วันที่สร้าง (Created date/time)
	วันที่เผยแพร่ (Released date/time)
แหล่งข้อมูล (Source)	ชื่อตำแหน่ง (Content location name)
	รหัสตำแหน่ง (Content location code)
	เมือง (City) สถานที่ (Location)
	เมือง/จังหวัด (State/Province)
	ประเทศ (Country)
	ชื่อวัตถุ (Object name)
	ตำแหน่งอ้างอิง (Original transmission reference)
	รหัสประเทศ (Country code)
รูปแบบอื่น ๆ (Options)	ตัวเลือกหลายรูปแบบ (Multiple selection)
	รูปแบบ (Mode)

การจดจำรูปภาพหรือการรู้จำวัตถุ ประกอบด้วยการตรวจหาวัตถุ การระบุภาพ และการจัดประเภทภาพ เป็นการระบุวัตถุที่น่าสนใจภายในรูปภาพและจำแนกว่าเป็นของประเภทใด การจดจำภาพด้วยปัญญาประดิษฐ์ เป็นปัญหาการวิจัยที่มีมายาวนานในด้านการมองเห็นด้วยคอมพิวเตอร์ ในขณะที่วิธีการต่าง ๆ พัฒนาขึ้นเมื่อเวลาผ่านไป ในปีที่ผ่านมาแมชชีนเลิร์นนิง (Machine Learning: ML) โดยเฉพาะอย่างยิ่งเทคโนโลยีการเรียนรู้เชิงลึก (deep learning) ประสบความสำเร็จอย่างมากในงานระบบการมองเห็นและการทำความเข้าใจภาพจำนวนมาก การจดจำรูปภาพด้วยแมชชีนเลิร์นนิงใช้อัลกอริทึมเพื่อเรียนรู้ความรู้ที่ซ่อนอยู่จากชุดข้อมูลของตัวอย่างที่ดีและไม่ดี (supervised learning) วิธีการเรียนรู้ของเครื่องที่ได้รับความนิยมมากที่สุดคือการเรียนรู้เชิงลึก โดยใช้เลเยอร์ที่ซ่อนอยู่หลายชั้นในแบบจำลอง ดังนั้นวิธีการรู้จำภาพการเรียนรู้เชิงลึกจึงให้ผลลัพธ์ที่ดีที่สุดในแง่ของประสิทธิภาพ และความยืดหยุ่น อย่างไรก็ตาม การเรียนรู้เชิงลึกจำเป็นต้องมีการติดฉลากข้อมูลด้วยตนเองเพื่อใส่คำอธิบายประกอบตัวอย่างที่ดีและไม่ดี การเรียนรู้ของแมชชีนนั้นเป็นไปในสองรูปแบบคือ การเรียนรู้โดยมีผู้สอน (Supervised) หรือการเรียนรู้โดยไม่มีผู้สอน (Unsupervised) การเรียนรู้โดยมีผู้สอนนั้นเครื่องจะเรียนรู้และทำนายผลลัพธ์ได้จากการช่วยเหลือของนักวิทยาศาสตร์ข้อมูล (Data Scientist) ส่วนการเรียนรู้โดยไม่มีผู้สอน (Unsupervised) นั้นเครื่องจะเรียนรู้และทำนายผลได้จากการจำแนกและสร้างแพทเทิร์นจากข้อมูลที่ได้รับ เมื่อเครื่องสามารถทำนายผลลัพธ์จากชุดข้อมูลจำนวนมากได้มากเท่าไร ก็จะยิ่งแสดงความสามารถในการเรียนรู้เชิงลึก (Deep Learning) มากเท่านั้น กระบวนการเรียนรู้จากข้อมูลที่มนุษย์กำกับไว้เรียกว่าการเรียนรู้แบบมีผู้สอน กระบวนการสร้างข้อมูลที่มีป้ายกำกับดังกล่าวเพื่อฝึกโมเดลปัญญาประดิษฐ์ ต้องใช้แรงงานที่ใช้เวลานาน และใช้ข้อมูลจำนวนมาก เพื่อให้เกิดประสิทธิภาพสูงสุด ตัวอย่างการรู้จำวัตถุในภาพถ่ายแสดงในภาพที่ 4

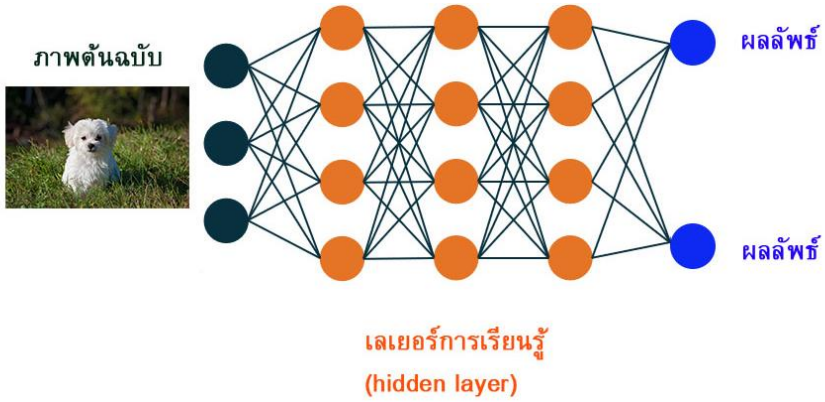
ภาพที่ 4 การรู้จำวัตถุในภาพถ่าย **(**ที่มา: https://aws.amazon.com/th/rekognition/)

ระบบเรียนรู้เชิงลึก (Deep learning: DL)

ปัญญาประดิษฐ์ (Artificial intelligence: AI) คือเทคโนโลยีที่สร้างให้คอมพิวเตอร์มีความฉลาดในลักษณะเดียวกับมนุษย์ แมชชีนเลิร์นนิง (Machine Learning: ML) คือเทคโนโลยีที่เกิดต่อมาจาก AI เป็นการทำให้ระบบคอมพิวเตอร์สามารถเรียนรู้ได้ด้วยตนเองโดยใช้ข้อมูล เป็นซับเซต (subset) ของ AI จุดประสงค์คือเพื่อใช้ในการสร้างแอปพลิเคชั่นที่มีประสิทธิภาพมากกว่ามนุษย์ โดยการทำให้ฉลาดขึ้น สามารถพัฒนา และเรียนรู้ได้ด้วยตนเองผ่านชุดข้อมูลที่สอน ส่วนการเรียนรู้เชิงลึก (Deep Learning: DL) เป็นซับเซตของแมชชีนเลิร์นนิง ที่สร้างโครงสร้างของอัลกอริธึมเป็นหลายเลเยอร์เพื่อสร้างโครงข่ายประสาทเทียมที่เรียนรู้และสร้างการตัดสินใจได้เอง มีวิธีการเรียนรู้ลักษณะต่างๆ ของข้อมูล ได้รับความนิยมในการจัดการแก้ปัญหาที่มีข้อมูลจำนวนมาก ซับซ้อน และไม่เป็นระบบ เช่นข้อมูลรูปภาพ

อัลกอริทึมการเรียนรู้ลึกเป็นเทคนิคที่นิยมในหลายสาขาด้านเทคโนโลยีเช่น การมองเห็นด้วยคอมพิวเตอร์ (computer vision), การประมวลผลภาษาธรรมชาติ (natural language processing), และการรู้จำเสียง (speech recognition) และอีกมากมาย หนึ่งในปัจจัยที่สำคัญที่สุดที่กำหนดประสิทธิภาพของอัลกอริทึมการเรียนรู้ลึกคืออัตราความแม่นยำ (accuracy rate) ซึ่งหมายถึงเปอร์เซ็นต์ของการจำแนกอย่างถูกต้องระหว่างจำนวนสิ่งทดสอบที่ถูกจัดอยู่ในชุดข้อมูลทั้งหมด

ภาพที่ 5 ความสัมพันธ์ระหว่างปัญญาประดิษฐ์ แมชชีนเลิร์นนิง และ การเรียนรู้เชิงลึก



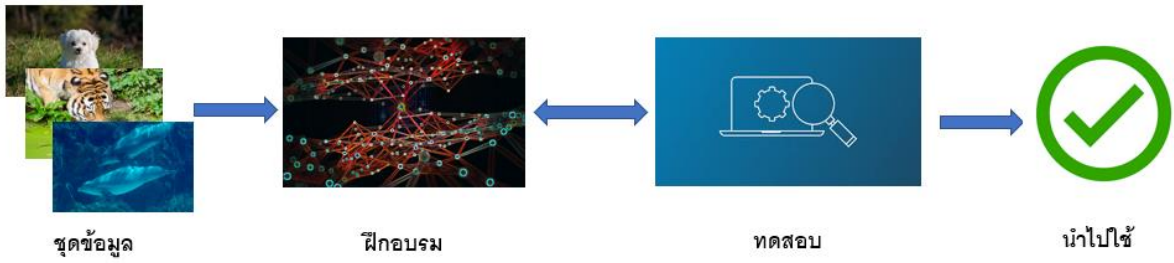
ภาพที่ 6 ขั้นตอนการทำงานของการเรียนรู้เชิงลึก

จากภาพที่ 6 มีการรับรูปภาพอินพุตเป็นรูปสุนัข รูปภาพนี้จะถูกจัดเก็บในรูปแบบของพิกเซลซึ่งใช้ 1 เลเยอร์ในการเก็บข้อมูลในรูปแบบสีขาวดำ (greyscale) และใช้ 3 เลเยอร์ ในการเก็บข้อมูลในรูปแบบรูสีต่างๆ มีการแยกคุณสมบัติ (feature) ที่สำคัญที่ใช้ในการจำแนกวัตถุในรูปภาพ การหาคุณสมบัติที่มีความเฉพาะเจาะจงต่อข้อมูลที่ได้รับมาจะทำที่เลเยอร์การเรียนรู้ (hidden layer) ยิ่งมีจำนวนเลเยอร์การเรียนรู้มาก โมเดลจะยิ่งมีความฉลาดมาก เมื่อทำการเรียนรู้เสร็จแล้ว โมเดลจะสามารถให้คำตอบว่าในแต่ละรูปภาพมีวัตถุชนิดใดอยู่บ้าง โดยพิจารณาจากวัตถุที่มีความน่าจะเป็นสูงสุด ก็จะให้ผลลัพธ์เป็นวัตถุสิ่งนั้น เช่น หากโมเดลวิเคราะห์ว่า รูปภาพมีความน่าจะเป็นสุนัขสูง ก็จะให้คำตอบว่ารูปภาพนั้นคือภาพสุนัข หากต้องการจำแนกให้ลึกซึ้งไปอีกว่าเป็นสุนัขสายพันธ์ใด สามารถทำได้โดยการเพิ่มคุณสมบัติของสายพันธ์สุนัขแต่ละชนิดในขั้นตอนการเรียนรู้ได้

กระบวนการของระบบจดจำภาพ (Image Recognition Process)

ขั้นตอนหลักในการทำงานของระบบจดจำภาพประกอบด้วยชุดข้อมูลพร้อมข้อมูลการฝึก การฝึกอบรมโครงข่ายประสาทเทียมเพื่อการจดจำภาพ และการทดสอบแบบจำลองปัญญาประดิษฐ์

1. **ชุดข้อมูลพร้อมข้อมูลการฝึก** โมเดลการรู้จำรูปภาพต้องการข้อมูลการฝึกโครงข่ายประสาทเทียมต้องการรูปภาพการฝึกอบรมจากชุดข้อมูลที่ได้รับเพื่อสร้างการรับรู้ว่าคุณลักษณะของวัตถุที่ตรวจพบว่าเป็นอย่างไร
2. **การฝึกอบรมโครงข่ายประสาทเทียมเพื่อการจดจำภาพ** ภาพจากชุดข้อมูลที่สร้างขึ้นจะถูกป้อนเข้าสู่อัลกอริธึมโครงข่ายประสาทเทียม นี่คือลักษณะการเรียนรู้เชิงลึกหรือแมชชีนเลิร์นนิงในการสร้างแบบจำลองการรู้จำรูปภาพ การฝึกอบรมอัลกอริธึมการรู้จำรูปภาพทำให้การจดจำรูปภาพของโครงข่ายประสาทเทียมสามารถระบุชนิดของวัตถุต่าง ๆ ได้
3. **การทดสอบแบบจำลองปัญญาประดิษฐ์** ต้องทดสอบโมเดลที่ได้รับการฝึกอบรมด้วยรูปภาพที่ไม่ได้เป็นส่วนหนึ่งของชุดข้อมูลการฝึก ใช้เพื่อกำหนดการใช้งาน ประสิทธิภาพ และความแม่นยำของโมเดล ดังนั้น ประมาณ 80-90% ของชุดข้อมูลรูปภาพทั้งหมดจะถูกใช้สำหรับการฝึกโมเดล ในขณะที่ข้อมูลที่เหลือจะถูกสงวนไว้สำหรับการทดสอบแบบจำลอง ประสิทธิภาพของโมเดลวัดจากชุดพารามิเตอร์ที่ระบุเปอร์เซ็นต์ความเชื่อมั่นของความถูกต้องต่อภาพทดสอบ



ภาพที่ 7 ขั้นตอนการจดจำภาพ

ซอฟต์แวร์จดจำรูปภาพ (Image Recognition Software)

ซอฟต์แวร์จดจำรูปภาพช่วยให้แอปพลิเคชันสามารถใช้อัลกอริธึมการเรียนรู้เชิงลึกเพื่อจดจำและเข้าใจรูปภาพหรือวิดีโอด้วยปัญญาประดิษฐ์ได้โดยผู้ใช้ไม่จำเป็นต้องมีความรู้ด้านการเขียนโปรแกรม ในปัจจุบันมีซอฟต์แวร์จดจำรูปภาพจำนวนมาก การเลือกใช้พิจารณาจาก ความสามารถในการรู้จำวัตถุ ติดตั้งและใช้งานง่าย ให้ประสิทธิภาพสูง มีการสอนโมเดลการรู้จำจากชุดข้อมูลจำนวนมาก มีเอพีไอ (API) และมีการพัฒนาอย่างต่อเนื่อง ในหัวข้อนี้จะนำเสนอโปรแกรมที่ใช้ในงานวิจัย 4 โปรแกรม ดังนี้

1. อเมซอน เรคคอคนิชัน (Amazon Rekognition) อเมซอน เรคคอคนิชัน เป็นชุดโปรแกรมที่พัฒนาโดยบริษัทอเมซอน เพื่อตรวจสอบ และวิเคราะห์ภาพอัตโนมัติด้วยปัญญาประดิษฐ์ (Addapa et al., 2020; Amazon Web Services, Inc., 2022) ผู้ใช้ไม่จำเป็นต้องมีความรู้ด้านปัญญาประดิษฐ์ก็สามารถใช้งานได้ มีโมเดลแมชชีนเลิร์นนิ่ง (Machine learning: ML) และคลังข้อมูลภาพขนาดใหญ่ เป็นแพลตฟอร์มประมวลผลบนระบบคลาวน์ (cloud) ด้วย อเมซอน เว็บ เซอร์วิส (Amazon Web Services: AWS) มีทั้งแบบให้ใช้งานแบบไม่มีค่าใช้จ่าย และเสียค่าใช้จ่ายหากต้องการคุณสมบัติที่เพิ่มมากขึ้น สามารถระบุวัตถุ บุคคล ข้อความ ฉาก และกิจกรรมในรูปภาพ วิเคราะห์ใบหน้าและการค้นหาใบหน้าที่แม่นยำสูงสามารถนำตัวกลางการเรียกใช้โปรแกรมหรือเอพีไอ (Application Program Interface: API) มาใช้งานเพื่อเชื่อมซอฟต์แวร์ภายนอกได้ อเมซอน เรคคอคนิชัน ใช้อัลกอริทึมการเรียนรู้เชิงลึก โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNNs) และ โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Networks: RNNs) ด้วยชุดข้อมูลกว่า 100,000 ภาพเพื่อฝึกโมเดลเรียนรู้ของเครื่อง สนับสนุนการตรวจจับวัตถุ, การรู้จำใบหน้า, การตรวจจับข้อความ, การรู้จำภาพทั่วไป, การรู้จำคนดัง, การตรวจสอบเนื้อหา, และป้ายกำกับที่กำหนดเอง การกำหนดราคาตามการใช้งานตามจำนวนรูปภาพที่ประมวลผลและระดับบริการที่ต้องการ ใช้ฟรี 1,000 รูปต่อเดือน สำหรับ 1 ล้านรูปต่อไป ราคา 0.001 ดอลล่าร์ต่อรูป หรือโดยประมาณราคา 1 ดอลล่าร์ต่อ 1,000 รูป

2. คลาริฟาย (Clarifai) คลาริฟาย (Clarifai, Inc., 2022) เป็นบริษัทที่พัฒนาโปรแกรมที่พัฒนาบนพื้นฐานของโครงข่ายประสาทเทียม ซึ่งเป็นเทคนิควิธีหนึ่งในปัญญาประดิษฐ์ทางด้านภาพถ่าย มีความสามารถในการจดใจวัตถุในภาพจากข้อมูลหรือแบบจำลองที่กำหนดเองได้ รวมถึงการค้นหาด้วยภาพ การรู้จำข้อความในภาพถ่าย และจัดระเบียบคอลเลคชันภาพถ่าย คลาริฟาย ใช้อัลกอริทึมการเรียนรู้เชิงลึก โครงข่ายความเชื่อแบบลึก (Deep Belief Networks: DBNs) โครงข่ายประสาทเทียมแบบคอนโวลูชัน และ โครงข่ายประสาทเทียมแบบวนกลับ ใช้ชุดข้อมูลภาพขนาดใหญ่ที่เป็นที่นิยมคืออิมเมจเน็ต (ImageNet) ซึ่งประกอบด้วยภาพกว่า 14 ล้านภาพที่แบ่งออกเป็นหมวดหมู่ต่างๆ ที่หลากหลายมากกว่า 21,000 หมวดหมู่ นอกจากนี้คลาริฟายยังใช้ชุดข้อมูลที่เป็นเจ้าของเองสำหรับแอปพลิเคชันบางประเภท เช่น แฟชั่นและอาหาร การกำหนดราคาตามการใช้งานตามจำนวนรูปภาพที่ประมวลผลและระดับบริการที่ต้องการ ใช้ฟรี 1,000 รูปต่อเดือน 5 แสนรูปต่อไปราคา 0.05 ดอลล่าร์ต่อรูป หรือโดยประมาณราคา 2 ดอลล่าร์ต่อ 1,000 รูป

3. อิมเมกกา (Imagga) อิมเมกกา (Imagga technologies, 2022) มีบริการหลักอยู่ 3 งานคือ การค้นหาความคล้ายคลึงของภาพ การติดแท็ก และการตรวจจับวัตถุในภาพ เป็นโปรแกรมที่ทำงานได้แบบครบถ้วน ทั้งการจำแนกประเภทวัตถุ การจัดหมวดหมู่ภาพ การกำหนดสี การจดจำใบหน้า การติดคำอธิบายภาพ การค้นหาภาพ ช่วยจัดระเบียบหมวดหมู่และการค้นหาภาพ ระบุส่วนประกอบของสิ่งของในภาพ อีกทั้งยังรองรับการแสดงผลเมทาดาทาได้หลายภาษา รวมถึงภาษาไทยด้วย อิมเมกกาใช้อัลกอริทึมการเรียนรู้เชิงลึก โครงข่ายประสาทเทียมแบบคอนโวลูชัน และ โครงข่ายประสาทเทียมแบบวนกลับ ด้วยชุดข้อมูลที่ของตัวเองที่มีภาพกว่า 20 ล้านภาพเพื่อฝึกโมเดลเรียนรู้ ชุดข้อมูลนี้ประกอบด้วยหมวดหมู่ภาพที่หลากหลาย เช่น คน สัตว์ พืช อาหาร และวัตถุ อีกทั้งยังมีชุดข้อมูลที่กำหนดเองสำหรับแอปพลิเคชันเฉพาะ เช่น แฟชั่นและอีคอมเมิร์ซ การกำหนดราคาตามการใช้บริการตามจำนวนรูปภาพที่ประมวลผลและระดับบริการที่ต้องการ โดยมีส่วนลดสำหรับการสมัครสมาชิกรายปี ใช้ฟรี 1,000 รูปต่อเดือน สำหรับ 70,000 รูปต่อไปราคาต่อเดือน 79 ดอลล่าร์ หรือโดยประมาณราคา 1.68 ดอลล่าร์ต่อ 1,000 รูป

4. กูเกิลคลาวน์วิชันเอพีไอ (Google cloud vision API) กูเกิลคลาวน์วิชันเอพีไอ (Google cloud platform, 2022) เป็นบริการของ Google Cloud Platform ที่ใช้ไลบรารี TensorFlow ในการพัฒนาโปรแกรมการวิเคราะห์ภาพ สามารถตรวจสอบวัตถุในภาพ อ่านข้อความในภาพ ตรวจสอบภาพไม่เหมาะสม ค้นหาใบหน้าคนภายในภาพ อารมณ์ของบุคคลในภาพ ตรวจสอบสถานที่ และรูปสัญลักษณ์ทางการค้าของผลิตภัณฑ์ในภาพ ข้อมูลเมทาดาทาที่ได้จากโปรแกรมเหล่านี้ จะประกอบด้วยข้อมูลคำและค่าความเชื่อมั่น (Confidence score) ของแต่ละคำ เพื่อให้ผู้ใช้งานได้ตัดสินใจเลือกใช้คำที่มีค่าความเชื่อมั่นที่ยอมรับได้ กูเกิลคลาวน์วิชันเอพีไอ ใช้อัลกอริทึมการเรียนรู้เชิงลึก โครงข่ายประสาทเทียมแบบคอนโวลูชัน ด้วยชุดข้อมูลระดับใหญ่ที่เรียกว่าโอเพนอิมเมจ (Open Images) ที่ประกอบด้วยภาพกว่า 9 ล้านภาพในหมวดหมู่กว่า 6,000 หมวดหมู่ ภาพในชุดข้อมูลมีป้ายกำกับด้วยคำอธิบายและหมวดหมู่ที่ช่วยให้ระบบจดจำรูปภาพได้ง่ายขึ้น กูเกิลยังมีชุดข้อมูลที่กำหนดเองสำหรับแอปพลิเคชันเฉพาะ เช่น การค้นหาในเว็บและอีคอมเมิร์ซ การกำหนดราคาตามการใช้บริการจ่ายตามจำนวนรูปภาพที่ประมวลผลและระดับของบริการที่ต้องการ โดยมีส่วนลดสำหรับการใช้บริการประจำปีหรือการใช้บริการจำนวนมาก ใช้ฟรี 1,000 รูปต่อเดือน 5 ล้านรูปต่อไปราคา 1.50 ดอลล่าร์ต่อ 1,000 รูป

หมายเหตุ เทคโนโลยี และชุดข้อมูล อาจมีการเปลี่ยนแปลงให้มีความถูกต้องแม่นยำขึ้น ตามเทคโนโลยีที่มีความก้าวหน้ามากขึ้น รวมถึงราคาค่าบริการที่อาจมีการเปลี่ยนแปลงได้ในอนาคต

เอพีไอ (Application Program Interface: API)

เอพีไอเป็นบริการช่องทางการเชื่อมต่อเพื่อแลกเปลี่ยนข้อมูลจากระบบหนึ่งไปสู่อีกระบบหนึ่ง ที่มีความสะดวก รวดเร็ว ปลอดภัย เป็นวิธีการเรียกใช้งานระหว่างโปรแกรมให้สามารถสื่อสารระหว่างกัน หน้าที่หลักของเอพีไอคือ คอยรับคำสั่งจากฝั่งลูกข่าย (Client) ซึ่งก็คือ แอปพลิเคชันต่าง ๆ เช่น เว็บแอปพลิเคชัน (Web Application) โมบายแอปพลิเคชัน (Mobile Application) เดสก์ท็อปแอปพลิเคชัน (Desktop Application) เป็นต้น เป็นช่องทางขอใช้บริการระหว่างโปรแกรม โดยโปรแกรมที่จะเรียกใช้งานต้องประกาศขึ้นมาก่อน เพื่อกำหนดรูปแบบการใช้งาน (Request/Response) โดยผู้ให้บริการ API จะเป็นผู้กำหนดคำสั่งในการเรียกใช้งาน ต้องมีการเชื่อมต่อกับอินเทอร์เนตเพื่อให้พร้อมเรียกใช้งานได้ทันที เมื่อฝั่งลูกข่าย (Client) ส่งคำสั่งจะเรียกว่าการรีเควส (Request) จากนั้นตัวเอพีไอจะรับคำสั่งไปประมวลผลและสรุปเป็นข้อมูลที่ตรงกับ Request และส่งข้อมูลเหล่านั้นกลับไปที่ฝั่งลูกข่าย หรือ แอปพลิเคชันเพื่อนำไปใช้งานต่อไป หรือกล่าวได้ว่าเอพีไอเป็นตัวกลางที่จะทำให้คอยรับคำสั่งต่าง ๆ ประมวลผลและกระทำข้อมูลส่งกลับคืนไปยังคนสั่งโดยอัตโนมัติโดยผู้เรียกใช้ไม่จำเป็นต้องรู้วิธีการพัฒนาหรือรายละเอียดภายในโปรแกรมซอฟต์แวร์นั้น

ภาพที่ 8 การทำงานของเอพีไอ

เอพีไอมีผลกระทบอย่างมากกับการพัฒนาซอฟต์แวร์ในปัจจุบันเพราะทำให้ซอฟต์แวร์ที่เขียนต่างภาษากัน ระบบปฏิบัติการต่างกันสามารถเรียกใช้งานกันได้ เช่นซอฟต์แวร์ที่สร้างโดยองค์กรที่ 1 สามารถเรียกใช้ซอฟต์แวร์ที่สร้างโดยองค์กรที่ 2 ซึ่งอาจมีความเชี่ยวชาญและทำผลิตภัณฑ์ที่ต่างกัน แต่เสริมกันได้เป็นอย่างดีสร้างประโยชน์ให้กับผู้ใช้อย่างมากด้วยลักษณะที่เป็นการแยกเป็นส่วน ๆ (modular) โปรแกรมเมอร์สามารถสร้างแอปพลิเคชันที่มีการทำงานต่างๆมากมายได้ด้วยการเรียกใช้เอพีไอของแต่ละบริการซอฟต์แวร์เฉพาะทางจากหลายที่ ทำให้ได้แอปพลิเคชันที่มีความสามารถสูงโดยที่โปรแกรมเมอร์ไม่จำเป็นต้องรู้สิ่งที่เกิดขึ้นภายในของซอฟต์แวร์ที่เรียกผ่านเอพีไอต่าง ๆ เลย API ทำงานใน 4 รูปแบบด้วยกัน ดังนี้

1. SOAP API ย่อมาจาก Simple Object Access Protocol เป็น API ที่ใช้โปรโตคอลการเข้าถึงอ็อบเจกต์อย่างง่าย ไคลเอ็นต์และเซิร์ฟเวอร์จะแลกเปลี่ยนข้อความโดยใช้ XML

2. RPC API ย่อมาจาก Remote Procedure Call เป็น API ที่เรียกใช้กระบวนการระยะไกล ไคลเอ็นต์ดำเนินการฟังก์ชัน (หรือกระบวนการ) หนึ่งๆ บนเซิร์ฟเวอร์ และเซิร์ฟเวอร์ส่งผลลัพธ์กลับไปยังไคลเอ็นต์

3. Websocket API คือการพัฒนา Web API สมัยใหม่ที่ใช้อ็อบเจกต์ JSON ในการส่งข้อมูล WebSocket API รองรับการสื่อสารสองทางระหว่างแอปไคลเอ็นต์และเซิร์ฟเวอร์ เซิร์ฟเวอร์สามารถส่งข้อความเรียกกลับไปยังไคลเอ็นต์ที่เชื่อมต่อ จึงทำให้มีประสิทธิภาพมากกว่า REST API

4. REST API ย่อมาจาก Representational State Transfer เป็น API ที่ได้รับความนิยมและยืดหยุ่นที่สุดที่พบในเว็บไซต์ปัจจุบัน ไคลเอ็นต์ส่งคำขอไปยังเซิร์ฟเวอร์เป็นข้อมูล เซิร์ฟเวอร์ใช้ข้อมูลอินพุตจากไคลเอ็นต์นี้เพื่อเริ่มต้นฟังก์ชันภายในและส่งคืนข้อมูลเอาต์พุตกลับไปยังไคลเอ็นต์ โดยอาศัยรูปแบบของ HTTP Method เช่น GET POST PUT DELETE ในการทำงาน และจะส่งค่ากลับมาเป็น json หรือ XML ส่งผลให้สามารถรับส่งข้อมูลไปมาข้ามแพลตฟอร์มได้อย่างสะดวก เพราะเป็นการเรียกผ่านโปรโตคอล HTTP ใช้ในการเรียกใช้เว็บไซต์ การทำงานของ rest นั้นจะอาศัย URI/URL ของ request เพื่อเป็นตัวค้นหาและประมวลผลแล้วตอบกลับโดย response

งานวิจัยที่เกี่ยวข้อง

ความสำคัญของการกำหนดเมทาดาทาแสดงในงานวิจัยของพิพิธภัณฑ์ธรรมศาสตร์เฉลิมพระเกียรติ คณะสังคมวิทยาและมานุษยวิทยา มหาวิทยาลัยธรรมศาสตร์ (ดุษฎีพร ชาติบุตร, 2562) ได้จัดทำโครงสร้างและรูปแบบเมทาดาทาข้อมูลดิจิทัลสำหรับพิพิธภัณฑ์ธรรมศาสตร์เฉลิมพระเกียรติ เพื่อให้เกิดมาตราฐานในการจัดการข้อมูลทางวัฒนธรรมในรูปแบบดิจิทัล การกำหนดเมทาดาทาที่เหมาะสมกับทรัพยากรสารสนเทศดิจิทัลมีความสำคัญยิ่งต่อการเข้าถึง และสืบค้น โดยเฉพาะทรัพยากรสารสนเทศที่มีลักษณะข้อมูลเฉพาะด้าน เช่นการพัฒนาเค้าร่างเมทาดาทาสำหรับจารึกที่อยู่ในรูปดิจิทัล (วิศปตย ชัยชวย , 2556) เมทาดาทาสำหรับภาพจิตรกรรมฝาผนัง (สุวรรณี ห้วยหงส์ทอง และกันยารัตน์ เควียเซ่น, 2562) เพื่อให้เกิดการจัดเก็บ และอนุรักษ์ภาพดิจิทัลอย่างมีประสิทธิภาพการประเมินคุณภาพของเมทาดาทาของสื่อในห้องสมุดดิจิทัล (Ochoa & Duval, 2009) โดยเปรียบเทียบการบันทึกเมทาดาทาแบบใช้มนุษย์ และแบบใช้เครื่องมืออัตโนมัติ พบว่ามนุษย์สามารถวิเคราะห์ส่วนประกอบที่ซับซ้อนของข้อมูลได้ดี ในขณะที่เครื่องมืออัตโนมัติสามารถเรียนรู้วัตถุที่อยู่ในภาพได้มากกว่า ในงานด้านบรรณารักษ์และหอสมุด สำนักหอสมุด มหาวิทยาลัยเชียงใหม่ ได้พัฒนาระบบเมทาดาทาเชิงสหสัมพันธ์ (ปราชญ์ สงวนศักดิ์, 2559) เพื่อเชื่อมโยงเมทาดาทาระบบห้องสมุดอัตโนมัติของห้องสมุดสถาบันอุดมศึกษาในประเทศไทยผ่านระบบคลาวน์ ทำให้เกิดการแบ่งปัน และใช้งานเมทาดาทาร่วมกัน งานวิจัยเรื่องการกำหนดเมทาดาทาสำหรับฐานข้อมูลดิจิทัลวัฒนธรรมและภูมิปัญญาท้องถิ่น (สิทธิชัย บวชไธสง, 2560) แสดงให้เห็นว่าการกำหนดเมทาดาทาที่เหมาะสม สามารถช่วยให้การจัดเก็บ และสืบค้นข้อมูล สะดวกรวดเร็วมีประสิทธิภาพมากยิ่งขึ้น

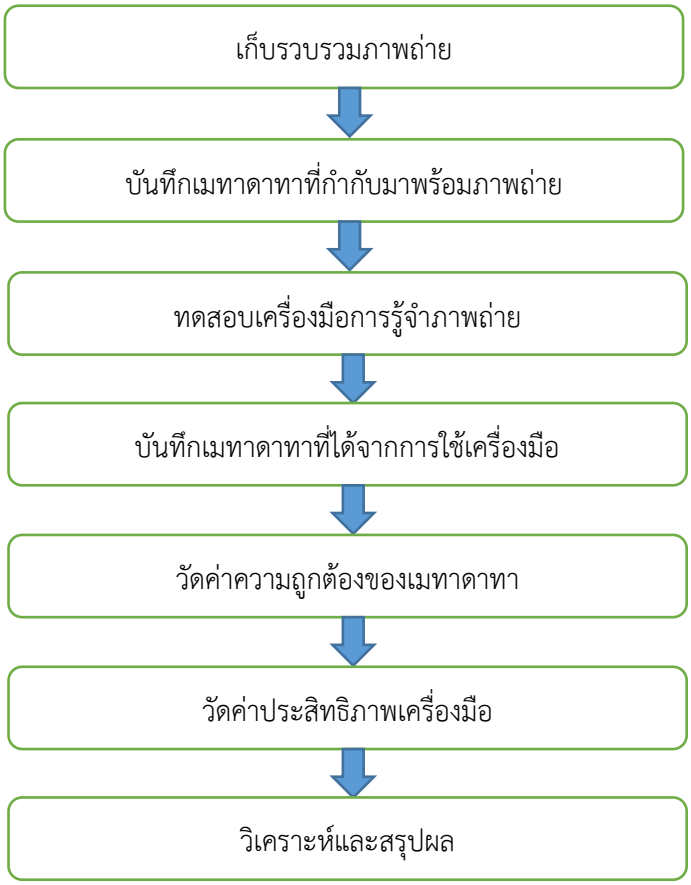
ข้อมูลเมทาดาทาสามารถนำไปใช้ได้ในอุตสาหกรรมหลายประเภท เช่น เว็บไซต์ที่ทำการค้าขายสินค้าที่มีรูปถ่ายจำนวนมาก เช่น ร้านขายเสื้อผ้า (Hong & Lee, 2015) เมื่อลูกค้าใช้คำค้นที่เฉพาะเจาะจงตามความต้องการของสีหรือเนื้อหา เว็บไซต์สามารถแสดงผลลัพธ์ได้อย่างแม่นยำโดยใช้ข้อมูลเมทาดาทาที่ใส่ไว้ในแต่ละรูปได้ เว็บไซต์ที่ให้บริการข้อมูลด้านศิลปวัฒนธรรม Europeana (Addapa et al., 2020) ที่มีการจัดการข้อมูลเมทาดาทาของภาพถ่าย เก็บข้อมูลเมทาดาทาอย่างเป็นระบบซึ่งช่วยให้การนำข้อมูลภาพถ่ายมาใช้สามารถเชื่อมโยงไปยังภาพอื่น ๆ ที่เกี่ยวข้อง ทำให้การค้นคว้าหาข้อมูลจากภาพถ่ายมีประสิทธิภาพ การตรวจสอบความถูกต้องของภาพถ่ายดิจิทัลด้วยข้อมูลเมทาดาทา ข้อมูล exif และการถอดรหัสข้อมูล (Harran et al., 2017; Gangwar & Pathania, 2018) การจำแนกภาพถ่ายด้วยข้อมูลเมทาดาทา exif (Ghazali et al. 2020) ด้วยอัลกอริทึมซัพพอร์ตเวคเตอร์แมชชีน (support vector machine) และการค้นหาข้อมูลที่ใกล้ที่สุด (K-nearest neighbor) มีค่าความถูกต้องที่ 75.15%

เทคโนโลยีปัญญาประดิษฐ์นำมาใช้ในการจัดการข้อมูลจดหมายเหตุ (Colavizza et al,. 2021) มากขึ้นเนื่องจากการเปลี่ยนแปลงรูปแบบการจัดเก็บแบบดั้งเดิมมาสู่คลังข้อมูลดิจิทัล เทคโนโลยีปัญญาประดิษฐ์มีบทบาทในหลายขั้นตอน เช่น การบันทึกข้อมูลและตัดสินใจแบบอัตโนมัติ การให้คำอธิบายและการดึงข้อมูลเมทาดาทาแบบอัตโนมัติ (Büttner, 2019) การสกัดเนื้อหาและการจัดทำดัชนีอัตโนมัติ (Binmakhashen, 2020) การรวมรูปแบบการจดจำภาพและข้อความเพื่อการรู้จำอักขระ (Muehlberger, 2019) การค้นหาและการจัดหมวดหมู่ข้อมูลที่เก็บในคลังดิจิทัลแบบอัตโนมัติ (Lee, 2019) งานวิจัยที่ใช้เมทาดาทาในภาพถ่าย ในการรู้จำหน้าบุคคล งานวิจัยที่ประเมินแอพพลิเคชันที่ใช้ปัญญาประดิษฐ์ในการจดจำใบหน้าในการเก็บภาพถ่ายดิจิทัลจากห้องสมุด พิพิธภัณฑ์ และสถาบันวัฒนธรรม (Bakker et al., 2020) โดยใช้ภาพถ่ายจากคอลเล็กชันดิจิทัลของมหาวิทยาลัยนานาชาติฟลอริดา (Florida International University) วิเคราะห์ด้วยแอพพลิชัน OpenCV, Face++ และ Amazon Rekognition ผลการวิจัยพบว่าแอพพลิเคชัน Amazon Rekognition ให้ผลลัพธ์ที่ดีที่สุด เนื่องจากมีคลังภาพที่ผ่านการสอนจำนวนมาก ทำให้การเรียนรู้ทำได้ได้ดีมากที่สุด แต่สิ่งที่ต้องคำนึงถึงมากขึ้นคือเรื่องลิขสิทธิ์ของบุคคลในภาพถ่าย การใช้เทคนิคนิวรอนเนตเวิร์คในการรู้จำใบหน้าและตรวจสอบตำแหน่งโดยใช้ข้อมูลเมทาดาทาในภาพถ่าย (Sushma & Kumari, 2021) การรู้จำภาพในการจัดหมวดหมู่ภาพถ่ายที่ได้จากโครงการ CoSoPho (Collecting Social Photography) ซึ่งเป็นภาพถ่ายที่เกิดจากกิจกรรมต่าง ๆ ในสังคม โดยวิเคราะห์ผลจากการใช้เครื่องมือ Amazon Rekognition, Clarifai และ Google Vision ผลการวิจัยใช้วิธีการให้คะแนนเมทาดาทาที่เครื่องมือแสดงผลเป็นช่วงคะแนนจากน้อยที่สุด (1) ถึง มากที่สุด (5) พบว่าการใช้เครื่องมือสามารถเพิ่มประสิทธิภาพการกำหนดเมทาดาทาได้ โดยเครื่องมือแต่ละตัวเหมาะกับการใช้งานในรูปภาพที่มีลักษณะที่แตกต่างกัน และยังไม่สามารถใช้เครื่องมือเพียงอย่างเดียวมาทดแทนมนุษย์ได้

บทที่ 3

วิธีทำการทดลอง

งานวิจัยนี้มุ่งนำเสนอการศึกษาเทคโนโลยีการรู้จำรูปภาพ ซึ่งเป็นเครื่องมือปัญญาประดิษฐ์แขนงหนึ่งมาใช้ในการกำหนดเมทาดาดาภาพถ่ายแบบอัตโนมัติ ซึ่งมีวิธีการดำเนินงาน โดยเริ่มต้นจากศึกษาลักษณะเมทาดาทาภาพถ่ายดิจิทัล ศึกษาเครื่องมือการรู้จำรูปภาพที่ใช้หลักการของปัญญาประดิษฐ์ ทำการทดลองรู้จำวัตถุในรูปภาพ จากนั้นนำผลการทดลองมาวิเคราะห์และเปรียบเทียบผลกับการกำหนดเมทาดาทาที่กำกับมาพร้อมภาพถ่ายเพื่อหาเครื่องมือที่เหมาะสมสำหรับรูปถ่ายแต่ละประเภท มีขั้นตอนวิธีการดำเนินการวิจัยดังแสดงในภาพที่ 9



ภาพที่ 9 ขั้นตอนการวิจัย

การวิจัยแบ่งเป็น 5 ขั้นตอนหลัก คือ การเก็บรวบรวมภาพถ่าย การบันทึกเมทาดาทาที่มาพร้อมภาพถ่าย การทดสอบเครื่องมือการรู้จำภาพถ่ายด้วยเทคนิคปัญญาประดิษฐ์ การบันทึกเมทาดาทาที่ได้จากเครื่องมือปัญญาประดิษฐ์ และการวิเคราะห์ประสิทธิภาพเมทาดาทาและเครื่องมือปัญญาประดิษฐ์ รายละเอียดของแต่ละขั้นตอนอธิบายได้ดังนี้

เก็บรวบรวมภาพถ่าย

ภาพถ่ายจากคลังภาพดิจิทัลแบบเปิด เผยแพร่ภายใต้โดเมนสาธารณะ ซึ่งไม่มีลิขสิทธิ์หรือเจ้าของลิขสิทธิ์ยกให้เป็นสาธารณะประโยชน์ จำนวน 1,000 ภาพ จากเว็บไซต์ครีเอทีฟคอมมอนส์ (https://creativecommons.org/) และเว็บไซต์ฟลิคเกอร์ (https://www.flickr.com/creativecommons/) โดยแบ่งหมวดหมู่ภาพถ่ายเป็น 10 หมวดหมู่ๆ ละ 100 ภาพ ประกอบด้วย

1. ภาพคน (human) ให้รหัสภาพถ่ายขึ้นต้นด้วย H
2. ภาพสัตว์ (animal) ให้รหัสภาพถ่ายขึ้นต้นด้วย A
3. ภาพต้นไม้ ดอกไม้ (plant and flower) ให้รหัสภาพถ่ายขึ้นต้นด้วย PF
4. ภาพวิวทิวทัศน์ (view and landscape) ให้รหัสภาพถ่ายขึ้นต้นด้วย VL
5. ภาพผัก ผลไม้ (vegetable and fruit) ให้รหัสภาพถ่ายขึ้นต้นด้วย VF
6. ภาพอาหาร (food) ให้รหัสภาพถ่ายขึ้นต้นด้วย F
7. ภาพยานพาหนะ (vehicle) ให้รหัสภาพถ่ายขึ้นต้นด้วย V
8. ภาพสถานที่ท่องเที่ยว (tourist landmark) ให้รหัสภาพถ่ายขึ้นต้นด้วย L
9. ภาพศิลปวัฒนธรรม (art and culture) ให้รหัสภาพถ่ายขึ้นต้นด้วย AC
10. ภาพปกหนังสือ/โปสเตอร์เก่า (old book cover and poster) ให้รหัสภาพถ่ายขึ้นต้นด้วย BP

โดยภาพถ่ายที่เก็บรวบรวมได้มีคุณสมบัติดังนี้

1. เป็นไฟล์รูปภาพประเภทไฟล์ .jpg หรือ .png
2. มีความละเอียดไม่น้อยกว่า 80 พิกเซล x 80 พิกเซล และไม่มากกว่า 4096 พิกเซล x 4096 พิกเซล
3. มีขนาดไฟล์ไม่มากกว่า 4 เมกะไบต์
4. มีเมทาดาทากำกับไม่น้อยกว่า 5 คำ และไม่มากกว่า 50 คำ

ตัวอย่างภาพถ่ายแต่ละหมวดหมู่แสดงใน ภาพที่ 10 ถึง ภาพที่ 19 ตามลำดับดังนี้

ภาพที่ 10 ตัวอย่างภาพถ่ายหมวดหมู่คน

ภาพที่ 11 ตัวอย่างภาพถ่ายหมวดหมู่สัตว์

ภาพที่ 12 ตัวอย่างภาพถ่ายหมวดหมู่ต้นไม้ ดอกไม้

ภาพที่ 13 ตัวอย่างภาพถ่ายหมวดหมู่วิวทิวทัศน์

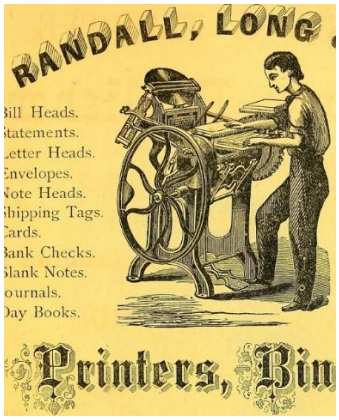
ภาพที่ 14 ตัวอย่างภาพถ่ายหมวดหมู่ผัก ผลไม้

ภาพที่ 15 ตัวอย่างภาพถ่ายหมวดหมู่อาหาร

ภาพที่ 16 ตัวอย่างภาพถ่ายหมวดหมู่ยานพาหนะ

ภาพที่ 17 ตัวอย่างภาพถ่ายหมวดหมู่สถานที่ท่องเที่ยว

ภาพที่ 18 ตัวอย่างภาพถ่ายหมวดหมู่ศิลปวัฒนธรรม



ภาพที่ 19 ตัวอย่างภาพถ่ายหมวดหมู่ปกหนังสือ/โปสเตอร์เก่า

บันทึกเมทาดาทาที่กำกับมาพร้อมภาพถ่าย

เครื่องมือที่ใช้ในการเก็บข้อมูล

แบบบันทึกข้อมูลเมทาดาทาที่กำกับมาพร้อมภาพถ่าย ประกอบด้วย รหัสไฟล์ภาพ (File ID) ชื่อภาพ (File name) ประเภทไฟล์ภาพ (File type) ขนาดไฟล์ (File size) ความละเอียดภาพ (Resolution) ลักษณะภาพ (Description) เจ้าของ (Author) แหล่งที่มา (Source) หมวดหมู่ (Category) สัญญาอนุญาต (License) ข้อมูลเมทาดาทาที่มาพร้อมภาพ (Metadata) ตัวอย่างภาพถ่ายรหัส H001 H002 และ A010

H001

PF001

A010

ภาพที่ 20 ตัวอย่างภาพถ่ายรหัส H001 PF001 และ A010

ตารางที่ 3 แบบบันทึกข้อมูลเมทาดาทาที่กำกับมาพร้อมภาพถ่าย

ลำดับที่ (No.)	รหัสไฟล์ภาพ (File ID)	ชื่อภาพ (File name)	ประเภทไฟล์ภาพ (File type)	ขนาดไฟล์ (File size)	ความละเอียดภาพ (Resolution:pixels)	ลักษณะภาพ (Description)	เจ้าของ (Author)	แหล่งที่มา (Source)	หมวดหมู่ (Category)	สัญญาอนุญาต (License)	เมทาดาทา (Metadata)
1	H001	Cycling couple	jpg	113 KB	800*450	ขี่จักรยาน	Pierre Blaché	https://www.flickr.com/photos/j0035004/51768477109/	Human	CC0	bike bicycle cycling people park garden
2	PF001	Fruit tree blossom	jpg	112 KB	400*300	ดอกไม้	Mathias Appel	https://www.flickr.com/photos/joelphick/29134130944/	Flower	CC0	Fruit Tree Blossom Flower Pint
3	A010	Giraffe Auckland Zoo	jpg	173 KB	800*598	ยีราฟ	Bernard Spragg. NZ	https://www.flickr.com/photos/volvob12b/9413151566/	Animal	CC0	Zoo Zebra Giraffe Creatures Animal
…											
1000											…

ทดสอบเครื่องมือการรู้จำภาพถ่ายด้วยเทคนิคปัญญาประดิษฐ์

ทดสอบเครื่องมือการรู้จำภาพถ่ายด้วยเทคนิคปัญญาประดิษฐ์จำนวน 4 โปรแกรมคือ

1. อเมซอน เรคคอคนิชัน (Amazon Rekognition) เข้าถึงได้จาก https://aws.amazon.com/rekognition/
2. คลาริฟาย (Clarifai) เข้าถึงได้จาก https://www.clarifai.com/
3. อิมเมกกา (Imagga) เข้าถึงได้จาก https://imagga.com/
4. กูเกิลคลาวน์วิชันเอพีไอ (Google cloud vision API) เข้าถึงได้จาก https://cloud.google.com/vision/

โปรแกรมทั้ง 4 โปรแกรมนี้ใช้เทคโนโลยีการเรียนรู้เชิงลึก (deep learning) ซึ่งเป็นวิทยาการแขนงหนึ่งในเรื่องปัญญาประดิษฐ์ที่ให้ผลลัพธ์ที่มีความถูกต้องสูง เพื่อวิเคราะห์ภาพ ตรวจจับวัตถุและฉากภายในรูปภาพ โดยผลลัพธ์ของทุกโปรแกรมจะประกอบด้วยข้อมูลเมทาดาทา และ ค่าระดับความเชื่อมั่น (Confidence levels) มี API ให้ใช้ทดลองใช้งานฟรี

ขั้นตอนการใช้งานเอพีไอ

1. สมัครลงทะเบียนผู้ใช้
2. เลือกลงทะเบียนเพื่อรับรหัสเอพีไอ (API Key) และรหัสผ่าน (API Secret)
3. ใช้งานด้วยคำสั่ง Requests

```
import requests
api_key = '&lt;replace-with-your-api-key&gt;'
api_secret = '&lt;replace-with-your-api-secret&gt;'
response = requests.get(
    'https://api.imagga.com/v2',
    auth=(api_key, api_secret))
print(response.json())
```

4. ได้ผลลัพธ์เมทาดาทา และค่าระดับความเชื่อมั่น

```
{Name: lighthouse, Confidence: 98.4629}
{Name: rock, Confidence: 79.2097}
{Name: sea, Confidence: 75.061}
```

อเมซอน เรคคอคนิชัน (Amazon Rekognition)

รองรับการทำงานได้หลายรูปแบบ ดังนี้

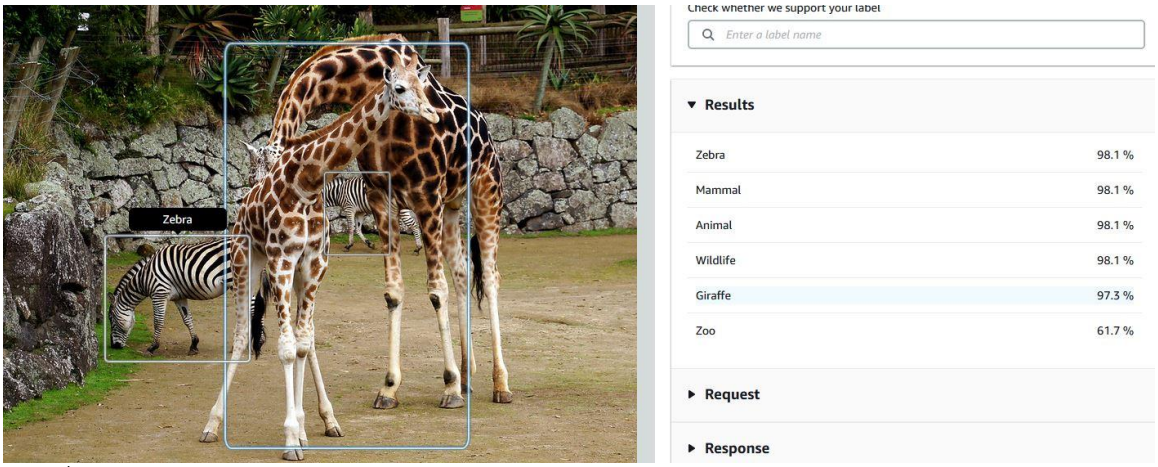
1. การระบุป้ายกำกับ (Labels) ระบุวัตถุ บุคคล ข้อความ ฉาก และกิจกรรมได้
2. การค้นหารูปภาพและวิดีโอ (Searchable image and video libraries) สามารถค้นหารูปภาพและวิดีโอที่จัดเก็บได้ สามารถค้นพบวัตถุและฉากที่ปรากฏขึ้นภายในได้
3. การตรวจสอบผู้ใช้ตามใบหน้า (Face-based user verification) ช่วยให้ยืนยันตัวตนผู้ใช้โดยเปรียบเทียบรูปภาพสดกับรูปภาพอ้างอิง
4. การตรวจจับอุปกรณ์ป้องกันภัยส่วนบุคคล (Detection of Personal Protective Equipment)ตรวจจับอุปกรณ์ป้องกันส่วนบุคคล (PPE) เช่น หน้ากาก ผ้าคลุมศีรษะ และผ้าคลุมมือ
5. การวิเคราะห์ความคิดเห็นและข้อมูลประชากร (Sentiment and demographic analysis) ตีความการแสดงออกทางอารมณ์ เช่น มีความสุข เศร้า หรือประหลาดใจ และข้อมูลทางประชากรศาสตร์ เช่น เพศจากภาพใบหน้า
6. การค้นหาใบหน้า (Facial Search) สามารถค้นหารูปภาพ วิดีโอที่จัดเก็บ และสตรีมวิดีโอสำหรับใบหน้าที่ตรงกับใบหน้าที่จัดเก็บไว้ในคอนเทนเนอร์ที่เรียกว่าคอลเลกชั่นใบหน้า
7. การตรวจจับเนื้อหาที่ไม่ปลอดภัย (Unsafe content detection) สามารถตรวจจับเนื้อหาสำหรับผู้ใหญ่และเนื้อหาที่มีความรุนแรงในรูปภาพและในวิดีโอที่จัดเก็บไว้
8. การจดจำผู้มีชื่อเสียง (Celebrity recognition) สามารถจดจำคนดังภายในรูปภาพและวิดีโอ
9. การตรวจจับข้อความ (Text detection) สามารถจดจำและแยกเนื้อหาที่เป็นข้อความออกจากรูปภาพได้
10. ระบุป้ายที่กำหนดเอง (Custom labels) สามารถระบุวัตถุและฉากในรูปภาพที่เจาะจงกับความต้องการได้

ขั้นตอนการทำงาน

ขั้นตอนที่ 1: ตั้งค่าบัญชี AWS และสร้างผู้ใช้ IAM
ขั้นตอนที่ 2: ตั้งค่า AWS CLI และ AWS SDK
ขั้นตอนที่ 3: เริ่มต้นใช้งาน AWS CLI และ AWS SDK API
ขั้นตอนที่ 4: เริ่มต้นใช้งานคอนโซล Amazon Rekognition

ใช้ AmazonRekognitionClassificationLabels_v2.0 ประกอบด้วยคำทั้งหมด 2580 คำ มี API ให้ใช้ทดลองใช้งานฟรี 5,000 ภาพต่อเดือน การดำเนินการของอเมซอน เรคคอคนิชัน เป็นแบบซิงโครนัส อินพุตและการตอบสนองอยู่ในรูปแบบ JSON จะวิเคราะห์รูปภาพอินพุตที่อยู่ในรูปแบบรูปภาพ

.jpg หรือ .png อีกทั้งยังได้รับการออกแบบมาให้ทำงานร่วมกับบริการของ AWS อื่นๆ เช่น Amazon S3 และ AWS Lambda ผลลัพธ์ที่ได้จะแสดงเมทาดาทา พร้อมค่าระดับความเชื่อมั่น พร้อมแสดงกรอบพื้นที่ของวัตถุที่ตรวจจับได้ ผลลัพธ์เมทาดาทาที่ได้จะเป็นภาษาอังกฤษ สามารถใช้ Amazon Translate เพื่อแปลเป็นภาษาอื่นได้ สามารถกรองและเลือกเฉพาะคำที่มีความมั่นใจสูง (เช่น 95% หรือสูงกว่า) หรือคำที่มีค่าความเชื่อมั่นต่ำกว่า (ใกล้ถึง 50%)



ภาพที่ 21 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมอเมซอน เรคคอคนิชัน

Request

```
{
    "Image": {
        "S3Object": {
            "Bucket": "MyBucket",
            "Name": "input.jpg"
        }
    },
    "MaxLabels": 10,
    "MinConfidence": 75
}
```

กำหนดจำนวนเมทาดาทา (MaxLabels) ไว้มากสุด 10 คำ และค่าระดับความเชื่อมั่นต่ำสุด (MinConfidence) ไว้ที่ 75%

การตรวจจับวัตถุ

```
#Copyright 2018 Amazon.com, Inc. or its affiliates. All Rights Reserved.
#PDX-License-Identifier: MIT-0 (For details, see https://github.com/awsdocs/amazon-rekognition-developer-guide/blob/master/LICENSE-SAMPLECODE.)
import boto3
def detect_labels(photo, bucket):
    client=boto3.client('rekognition')
    response = client.detect_labels(Image={'S3Object':{'Bucket':bucket,'Name':photo}}, MaxLabels=10)
    print('Detected labels for ' + photo)
    for label in response['Labels']:
        print ("Label: " + label['Name'])
        print ("Confidence: " + str(label['Confidence']))
        print ("Instances:")
        for instance in label['Instances']:
            print ("  Bounding box")
            print ("    Top: " + str(instance['BoundingBox']['Top']))
            print ("    Left: " + str(instance['BoundingBox']['Left']))
            print ("    Width: " +  str(instance['BoundingBox']['Width']))
            print ("    Height: " +  str(instance['BoundingBox']['Height']))
            print ("  Confidence: " + str(instance['Confidence']))
            print()
        print ("Parents:")
        for parent in label['Parents']:
            print ("   " + parent['Name'])
        print ("----------")
    return len(response['Labels'])
def main():
    photo=''
    bucket=''
    label_count=detect_labels(photo, bucket)
    print("Labels detected: " + str(label_count))
if __name__ == "__main__":
    main()
```

Response


```
{    "Labels": [
        {   "Name": "Zebra",
            "Confidence": 98.19500732421875,
            "Instances": [
                {
                    "BoundingBox": {
                        "Width": 0.21801050007343292,
                        "Height": 0.25438207387924194,
                        "Left": 0.18564695119857788,
                        "Top": 0.5029727816581726
                    },
                    "Confidence": 98.19500732421875
                },
                {
                    "BoundingBox": {
                        "Width": 0.09933410584926605,
                        "Height": 0.16792772710323334,
                        "Left": 0.5208384990692139,
                        "Top": 0.37427759170532227
                    },
                    "Confidence": 63.30255889892578
                }
            ],
            "Parents": [
                {
                    "Name": "Wildlife"
                },
                {
                    "Name": "Mammal"
                },
                {
                    "Name": "Animal"
```

```
            }
        ]
    },
    {
        "Name": "Mammal",
        "Confidence": 98.19500732421875,
        "Instances": [],
        "Parents": [
            {
                "Name": "Animal"
            }
        ]
    },
    {
        "Name": "Animal",
        "Confidence": 98.19500732421875,
        "Instances": [],
        "Parents": []
    },
    {
        "Name": "Wildlife",
        "Confidence": 98.19500732421875,
        "Instances": [],
        "Parents": [
            {
                "Name": "Animal"
            }
        ]
    }, {
        "Name": "Giraffe",
        "Confidence": 97.3818359375,
        "Instances": [
            {
```

```
                "BoundingBox": {
                    "Width": 0.36936891078948975,
                    "Height": 0.8236808776855469,
                    "Left": 0.36908936500549316,
                    "Top": 0.11070603877305984
                },
                "Confidence": 97.3818359375
            }
        ],
        "Parents": [
            {
                "Name": "Wildlife"
            },
            {
                "Name": "Mammal"
            },
            {
                "Name": "Animal"
            }
        ]
    },
    {
        "Name": "Zoo",
        "Confidence": 61.739158630371094,
        "Instances": [],
        "Parents": [
            {
                "Name": "Animal"
            }
        ]
    }
],
"LabelModelVersion": "2.0" }
```

แสดงเมทาดาทาไว้ในตัวแปร Name และค่าระดับความเชื่อมั่นไว้ในตัวแปร Confidence ดังนี้

"Name": "Zebra",

"Confidence": 98.19500732421875,

หมายถึง เมทาดาทาคือคำว่า Zebra มีค่าระดับความเชื่อมั่นที่ 98.19%

ระบุตำแหน่งขอบเขตของกล่องวัตถุที่ตรวจพบด้วยตัวแปร BoundingBox ระบุค่าความกว้าง (width) ความสูง (height) ตำแหน่งด้านซ้าย (left) และตำแหน่งด้านบน (top)

"BoundingBox": {

"Width": 0.21801050007343292,

"Height": 0.25438207387924194,

"Left": 0.18564695119857788,

"Top": 0.5029727816581726

},

"Confidence": 98.19500732421875

คลาริฟาย (Clarifai)

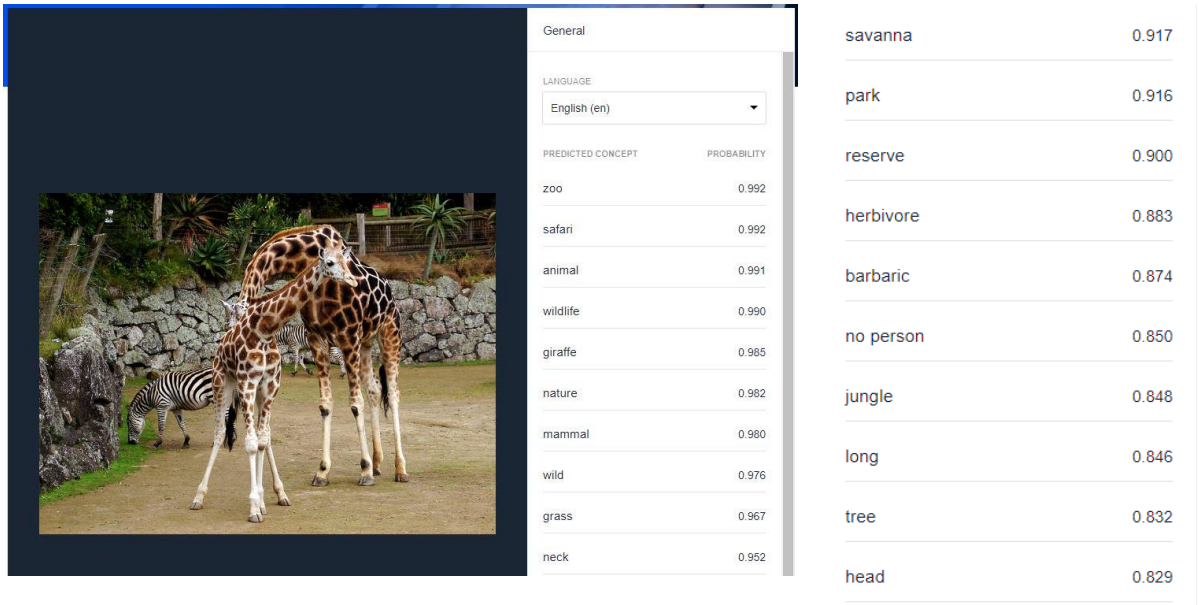
โปรแกรมนี้ใช้เทคนิค RetinaNet พร้อมสถาปัตยกรรมแกนหลัก Inception-V4 โดยฝึกอบรมโมเดลด้วยชุดข้อมูล Open Images V4 เป็นหลัก ซึ่งชุดข้อมูลนี้ประกอบด้วยวัตถุที่มีป้ายกำกับมากกว่า 30 ล้านรายการ สามารถส่งภาพได้สูงสุด 128 ภาพในการเรียก API ครั้งเดียว

รองรับการทำงานได้หลายรูปแบบ ดังนี้

1. การดูแลข้อความ (Text moderation) ตรวจจับภาษาที่ไม่เหมาะสม ลามกอนาจาร คุกคาม
2. การฝังข้อความ (Text embedding) การกำหนดแท็กหรือหมวดหมู่เพื่อวิเคราะห์ข้อความตามเนื้อหา
3. การดูแลภาพ (Image moderation) ระบุเนื้อหาที่ไม่ต้องการ มีความรุนแรง
4. ทั่วไป (General) แสดงป้ายกำกับ
5. กำหนดเอง (Custom) สร้างแบบจำลองเฉพาะเรื่องที่ต้องการ
6. ข้อมูลประชากร (Demographics) ทำนายอายุ เพศ เชื้อชาติจากภาพถ่าย
7. ตรวจจับใบหน้า (Face Detection)
8. ตรวจจับภาพคนดัง (Celebrity)
9. ตรวจจับเครื่องแต่งกาย (Apparel Detection)
10. ตรวจจับอาหารและส่วนผสม (Food)
11. ตรวจจับสี (Color)

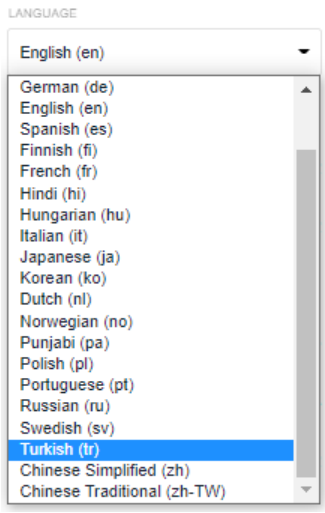
12. ตรวจจับพื้นผิวและลวดลาย (Textures and patterns)
13. ตรวจจับภาพลามก (Not safe for work)
14. ตรวจจับภาพเกี่ยวกับการท่องเที่ยว (Travel)
15. ตรวจจับภาพเกี่ยวกับการแต่งงาน (Wedding)

แสดงผลเมทาดาทา 20 อันดับแรก ซึ่งค่าความเชื่อมั่นจะไม่น้อยกว่า 0.70 หรือ 70%



ภาพที่ 22 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมคลาริฟาย

มี API ให้ใช้ทดลองใช้งานฟรี 1,000 ภาพต่อเดือน รองรับภาษา 53 ภาษา แต่ไม่รองรับภาษาไทย รองรับไฟล์ภาพนามสกุล .jpg .png .tiff .bmp และ .webp



Python-Request

```
from clarifai_grpc.grpc.api import service_pb2, resources_pb2
from clarifai_grpc.grpc.api.status import status_code_pb2

YOUR_CLARIFAI_API_KEY = "???"
YOUR_APPLICATION_ID = "???"
SAMPLE_URL = "https://samples.clarifai.com/metro-north.jpg"

# This is how you authenticate.
metadata = (("authorization", f"Key {YOUR_CLARIFAI_API_KEY}"),)

request = service_pb2.PostModelOutputsRequest(
    # This is the model ID of a publicly available General model. You may use any other
public or custom model ID.
    model_id="general-image-recognition",
    user_app_id=resources_pb2.UserAppIDSet(app_id=YOUR_APPLICATION_ID),
    inputs=[
        resources_pb2.Input(
            data=resources_pb2.Data(image=resources_pb2.Image(url=SAMPLE_URL))
        )
    ],
)
response = stub.PostModelOutputs(request, metadata=metadata)

if response.status.code != status_code_pb2.SUCCESS:
    print(response)
    raise Exception(f"Request failed, status code: {response.status}")

for concept in response.outputs[0].data.concepts:
    print("%12s: %.2f" % (concept.name, concept.value))
```

Response

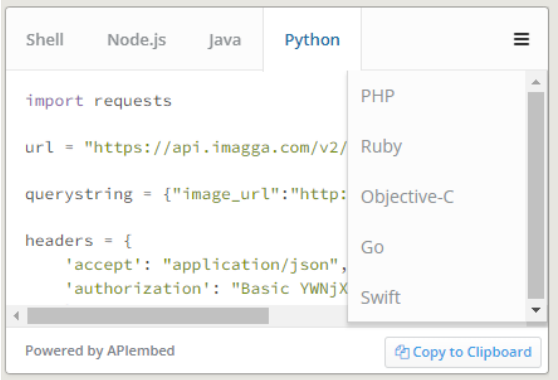
```
data {
  concepts {
    id: "ai_HLmqFqBf"
    name: "train"
    value: 0.9996053576469421
    app_id: "main"
  }
  concepts {
    id: "ai_fvlBqXZR"
    name: "railway"
    value: 0.9992986917495728
    app_id: "main"
  }
```

อิมเมกกา (Imagga)

รองรับการทำงานได้หลายรูปแบบ ดังนี้

1. การอธิบายภาพ (Tag) ใส่คำอธิบายประกอบภาพ
2. การจัดหมวดหมู่ (Categorization) จัดเรียงรูปภาพเป็นหมวดหมู่ที่เกี่ยวข้อง
3. การกลั่นกรองเนื้อหาของภาพ (Content moderation) กลั่นกรองเนื้อหารูปภาพสำหรับผู้ใหญ่
4. การครอบตัด (Cropping) เลือกพื้นที่ที่ดึงดูดสายตามากที่สุดของรูปภาพ
5. การแยกสี (Color extraction) ค้นหาสีที่โดดเด่นของวัตถุพื้นหลังและพื้นหน้าของรูปภาพ
6. การจดจำใบหน้า (Face recognition) ตรวจจับและจดจำภาพของผู้คน
7. การค้นหาด้วยภาพ (Visual search) อนุญาตให้ค้นหาผ่านเนื้อหาที่คล้ายคลึงกัน
8. การฝึกอบรมแบบกำหนดเอง (Custom training) จัดหมวดหมู่ให้เหมาะสมกับเนื้อหาเฉพาะ

มี API ให้ใช้ทดลองใช้งานฟรี 1,000 ภาพต่อเดือน ผลลัพธ์ที่ได้แสดงผลในรูปแบบของอ็อบเจ็กต์ JSON (JavaScript Object Notation) ซึ่งมี 2 ส่วนหลักประกอบด้วย แท็ก - อาร์เรย์ของแท็กทั้งหมดที่แนะนำสำหรับแต่ละภาพ และค่าระดับความเชื่อมั่นเป็นตัวเลขที่แสดงเปอร์เซ็นต์จาก 0 ถึง 100 โดยที่ 100 หมายความว่าแท็กนี้มีความเกี่ยวข้องสูง และความเชื่อมั่นน้อยกว่า 30 หมายความว่าแท็กนี้มีความเกี่ยวข้องต่ำ โดยกำหนดค่าการแสดงผลเริ่มต้นที่ระดับความเชื่อมั่น 10 อันดับแรก และสามารถดูผลเพิ่มเติมที่ค่าระดับความเชื่อมั่นที่ต่ำกว่าไปจนถึง 7% ได้ ผลลัพธ์รองรับภาษาต่างๆ ถึง 50 ภาษารวมถึงภาษาไทยด้วย



Requests

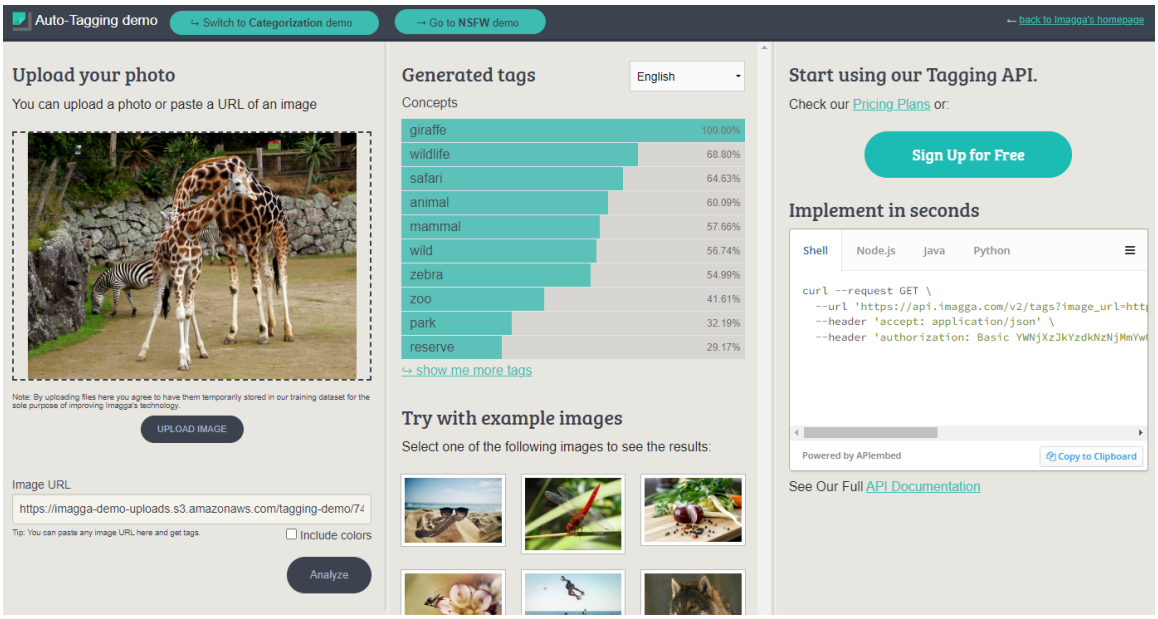
```
import requests
url = "https://api.imagga.com/v2/tags"
querystring = {"image_url":"http://playground.imagga.com/static/img/example_photo.jpg","version":"2"}
headers = {
    'accept': "application/json",
    'authorization': "Basic YWNjXzJkYzdkNzNjMmYwODliMToxYzQ3Yzg2ZDg0YjdmYjdjYjZjNzQ1NTQ1MmYwNTgzMQ==" }
response = requests.request("GET", url, headers=headers, params=querystring)
print(response.text)
```

API-Python

```
import requests
api_key = '&lt;replace-with-your-api-key&gt;'
api_secret = '&lt;replace-with-your-api-secret&gt;'
image_path = '/path/to/your/image.jpg'
response = requests.post(
    'https://api.imagga.com/v2/tags',
    auth=(api_key, api_secret),
    files={'image': open(image_path, 'rb')})
print(response.json())
```

Response


```
{
  "result": {
    "tags": [
      {
        "confidence": 61.4116096496582,
        "tag": {
          "en": "mountain"
        }
      },
      {
        "confidence": 54.3507270812988,
        "tag": {
          "en": "landscape"
        }
      },
```

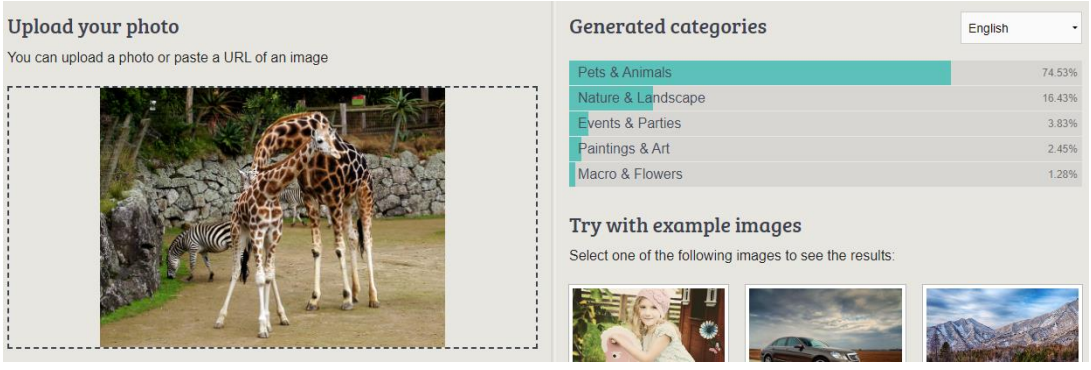
ภาพที่ 23 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมอิมเมกกา

ภาพที่ 24 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมอิมเมกกาแบบยาวเป็นภาษาอังกฤษ

ภาพที่ 25 ตัวอย่างผลการแยกหมวดหมู่ภาพถ่ายด้วยโปรแกรมอิมเมกกาเป็นภาษาไทย

โปรแกรมอิมเมกกายังสามารถแยกหมวดหมู่ของภาพได้ทั้งหมด 13 หมวดหมู่ คือ

1. ภายในอาคาร วัตถุ (interior and objects)
2. ภูมิทัศน์ (nature and landscape)
3. ชายหาด ชายทะเล (beaches and seaside)
4. งานปาร์ตี้ (events and parties)
5. อาหาร เครื่องดื่ม (food and drinks)
6. ภาพวาด งานศิลปะ (paintings and art)
7. สัตว์เลี้ยง สัตว์ป่า (pets and animals)
8. ข้อความ ภาพ (text and visuals)
9. พระอาทิตย์ขึ้น พระอาทิตย์ตก (sunrises and sunsets)
10. รถยนต์ ยานพาหนะ (cars and vehicles)
11. แมโคร ดอกไม้ (macro and flowers)
12. มุมมองถนน สถาปัตยกรรม (street view and architecture)
13. คน (people and portraits)



ภาพที่ 26 ตัวอย่างการแยกหมวดหมู่ด้วยโปรแกรมอิมเมกกา

กูเกิลคลาวน์วิชันเอพีไอ (Google cloud vision API)

กูเกิลคลาวน์วิชันเอพีไอ เป็นหนึ่งในบริการของกูเกิลคลาวแพลตฟอร์ม (Google Cloud Platform) ที่ใช้ความสามารถจากเทคโนโลยีของกูเกิล เพื่อช่วยอำนวยความสะดวกในเรื่องการวิเคราะห์ภาพ ซึ่งเอพีไอตัวนี้จะช่วยสามารถเข้าใจได้ว่าเนื้อหาของภาพนั้นคืออะไร โดยใช้เทคโนโลยีตัวเดียวกันกับที่ใช้ในกูเกิลโฟโต้ (Google Photos) ด้วย สามารถวิเคราะห์ภาพถ่ายได้อย่างแม่นยำที่สามารถนำไปใช้งานได้ทันทีโดยไม่ต้องกังวลถึงขั้นตอนการพัฒนา ติดตั้ง และคุณภาพของโมเดล เพราะกูเกิลจะเป็นผู้จัดการพัฒนาและปรับปรุงประสิทธิภาพอยู่เบื้องหลังทั้งหมด โดยใช้เทคโนโลยีการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning :RL) สามารถจัดการกับโมเดลทั้งหมดได้ผ่านหน้าคอนโซล (console) ของกูเกิลโดยตรง สามารถสร้างโมเดลขึ้นจากชุดข้อมูลของผู้ใช้ได้เอง เทคโนโลยีการเรียนรู้แบบเสริมกำลังที่ประกอบไปด้วยนิวรอนเนตเวิร์ค (neural network) ตัวควบคุม (controller) ที่สามารถเสนอไอเดียกลไกการทำงานของโมเดล และระบบที่คอยทดสอบและให้ผลตอบกลับ (feedback) ว่าตัวควบคุมควรปรับปรุงโมเดลที่เสนอมานี้อย่างไรบ้าง หลังการเสนอไอเดีย ทดสอบ และปรับปรุง ท้ายที่สุดแล้วตัวควบคุมจะเสนอโมเดลที่สามารถทำงานได้อย่างดีที่สุด และยังมีการเรียกใช้งาน TensorFlow ซึ่งเป็นไลบรารีสำหรับการเรียนรู้เชิงลึก ที่ใช้เพิ่มประสิทธิภาพกับงานต่างๆเช่น เครื่องมือค้นหา (search engine) การแปลภาษา (translation) คำบรรยายภาพ (image captioning) และ เครื่องมือช่วยการเสนอแนะ (recommendations) มี API ให้ใช้ทดลองใช้งานฟรี 1,000 ภาพต่อเดือน

รองรับการทำงานได้หลายรูปแบบ ดังนี้

1. กำหนดจุดที่แนะนำสำหรับพื้นที่ครอบตัดบนรูปภาพ (CROP_HINTS)
2. ตรวจจับข้อความในเอกสาร (DOCUMENT_TEXT_DETECTION)
3. ตรวจจับใบหน้าภายในภาพ (FACE_DETECTION)
4. คำนวณชุดคุณสมบัติของรูปภาพ เช่น สีที่โดดเด่นของรูปภาพ (IMAGE_PROPERTIES)
5. ตรวจจับป้ายกำกับตามเนื้อหารูปภาพ (LABEL_DETECTION)
6. ตรวจจับจุดสังเกตทางภูมิศาสตร์ภายในภาพ (LANDMARK_DETECTION)
7. ตรวจจับโลโก้บริษัทภายในภาพ (LOGO_DETECTION)
8. ตรวจหาเนื้อหาที่อาจไม่ปลอดภัยหรือไม่พึงประสงค์ (SAFE_SEARCH_DETECTION)
9. ตรวจจับข้อความในรูปภาพ (TEXT_DETECTION)
10. ตรวจจับบุคคลเฉพาะเรื่อง (WEB_DETECTION)

ขั้นตอนการทำงาน

1. ลงชื่อเข้าใช้บัญชี Google Cloud
2. ในคอนโซล Google Cloud ในหน้าตัวเลือกโปรเจ็กต์ เลือกหรือสร้างโปรเจ็กต์ Google Cloud
3. เปิดใช้งาน Cloud Vision API
4. สร้างคีย์เอพีไอ และรหัสผ่าน
5. ตั้งค่าตัวแปรสภาพแวดล้อม GOOGLE_APPLICATION_CREDENTIALS เป็นเส้นทางของไฟล์ JSON

```
def detect_labels(path):
    """Detects labels in the file."""
    from google.cloud import vision
    import io
    client = vision.ImageAnnotatorClient()
    with io.open(path, 'rb') as image_file:
        content = image_file.read()

    image = vision.Image(content=content)
    response = client.label_detection(image=image)
    labels = response.label_annotations
    print('Labels:')

    for label in labels:
        print(label.description)

    if response.error.message:
        raise Exception(
            '{}\nFor more info on error messages, check: '
            'https://cloud.google.com/apis/design/errors'.format(
                response.error.message))
```

Request

```
{
  "requests": [
    {
      "features": [
        {
          "maxResults": 50,
          "type": "LANDMARK_DETECTION"
        },
        {
          "maxResults": 50,
          "type": "FACE_DETECTION"
        },
        {
          "maxResults": 50,
          "type": "OBJECT_LOCALIZATION"
        },
        {
          "maxResults": 50,
          "type": "LOGO_DETECTION"
        },
        {
          "maxResults": 50,
          "type": "LABEL_DETECTION"
        },
        {
          "maxResults": 50,
          "model": "builtin/latest",
          "type": "DOCUMENT_TEXT_DETECTION"
        },
        {
          "maxResults": 50,
          "type": "SAFE_SEARCH_DETECTION"
```

```
        },
        {
          "maxResults": 50,
          "type": "IMAGE_PROPERTIES"
        },
        {
          "maxResults": 50,
          "type": "CROP_HINTS"
        }
      ],
      "image": {
        "content": "(data from A10.jpg)"
      },
      "imageContext": {
        "cropHintsParams": {
          "aspectRatios": [
            0.8,
            1,
            1.2
          ]
        }
      }
    }
  ]
}
```


Response


```
{
  "labelAnnotations": [
    {
      "description": "Giraffidae",
      "mid": "/m/01v46r",
      "score": 0.92908275,
```

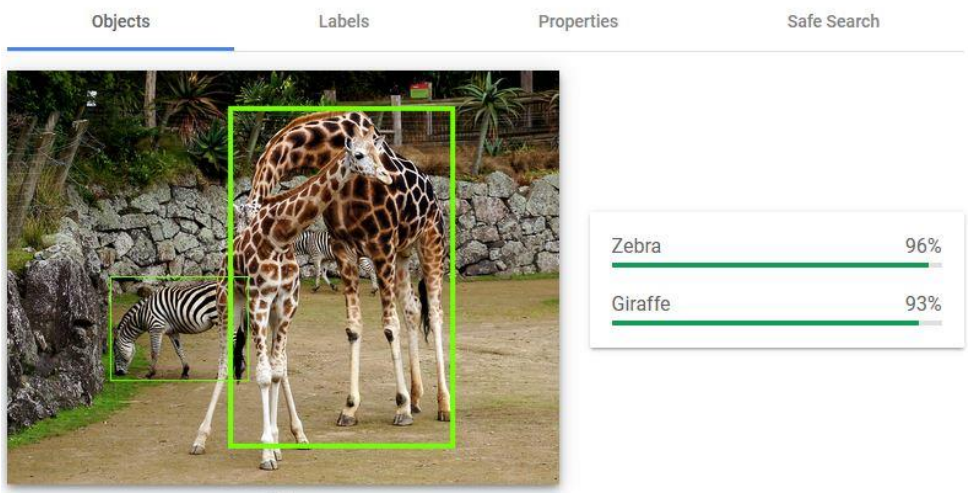
```
  "topicality": 0.92908275
},
{
  "description": "Vertebrate",
  "mid": "/m/09686",
  "score": 0.919814,
  "topicality": 0.919814
},
{
  "description": "Plant",
  "mid": "/m/05s2s",
  "score": 0.91838485,
  "topicality": 0.91838485
},
{
  "description": "Giraffe",
  "mid": "/m/03bk1",
  "score": 0.91128004,
  "topicality": 0.91128004
},
{
  "description": "Neck",
  "mid": "/m/0dzd8",
  "score": 0.88005626,
  "topicality": 0.88005626
},
{
  "description": "Organism",
  "mid": "/m/05nnm",
  "score": 0.867331,
  "topicality": 0.867331
},
{
```

```
  "description": "Fawn",
  "mid": "/m/0276krm",
  "score": 0.81594026,
  "topicality": 0.81594026
},
{
  "description": "Terrestrial animal",
  "mid": "/m/0fbf1m",
  "score": 0.8136724,
  "topicality": 0.8136724
},
{
  "description": "Grass",
  "mid": "/m/08t9c_",
  "score": 0.7986731,
  "topicality": 0.7986731
},
{
  "description": "Zebra",
  "mid": "/m/0898b",
  "score": 0.7889395,
  "topicality": 0.7889395
},
{
  "description": "Landscape",
  "mid": "/m/025s3q0",
  "score": 0.7754316,
  "topicality": 0.7754316
},
{
  "description": "Snout",
  "mid": "/m/05mqq3",
  "score": 0.771818,
```

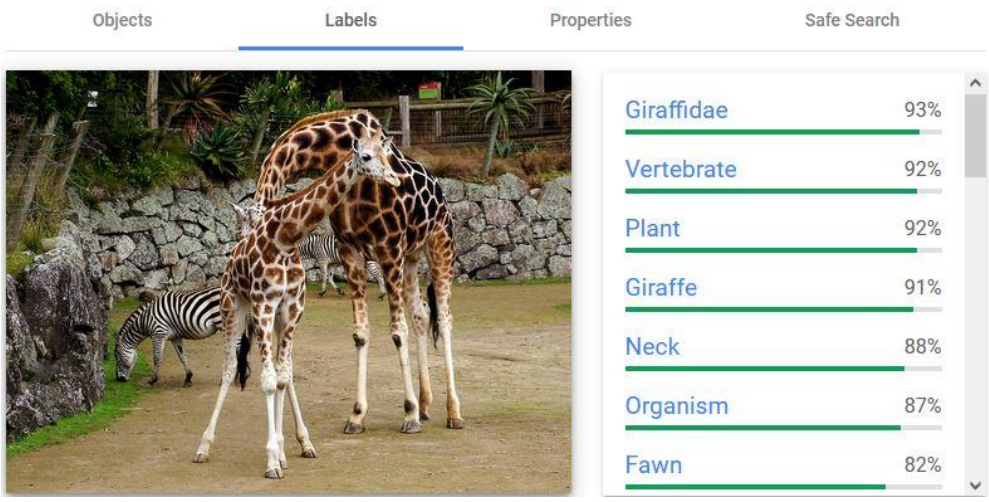
```
  "topicality": 0.771818
},
{
  "description": "Grassland",
  "mid": "/m/01c7cq",
  "score": 0.7633734,
  "topicality": 0.7633734
},
{
  "description": "Recreation",
  "mid": "/m/06bm2",
  "score": 0.75889146,
  "topicality": 0.75889146
},
{
  "description": "Wildlife",
  "mid": "/m/01280g",
  "score": 0.7155778,
  "topicality": 0.7155778
},
{
  "description": "Nature reserve",
  "mid": "/m/0swq_",
  "score": 0.6671671,
  "topicality": 0.6671671
},
{
  "description": "Pattern",
  "mid": "/m/0hwky",
  "score": 0.6349945,
  "topicality": 0.6349945
},
{
```

```
  "description": "Jungle",
  "mid": "/m/01y3fy",
  "score": 0.6068363,
  "topicality": 0.6068363
},
{
  "description": "Tree",
  "mid": "/m/07j7r",
  "score": 0.58128315,
  "topicality": 0.58128315
},
{
  "description": "Palm tree",
  "mid": "/m/0cdl1",
  "score": 0.5688465,
  "topicality": 0.5688465
},
{
  "description": "Tail",
  "mid": "/m/06z_nw",
  "score": 0.5577832,
  "topicality": 0.5577832
},
{
  "description": "Zoo",
  "mid": "/m/089v3",
  "score": 0.5489116,
  "topicality": 0.5489116
},
{
  "description": "Trunk",
  "mid": "/m/02tcwp",
  "score": 0.5251859,
```

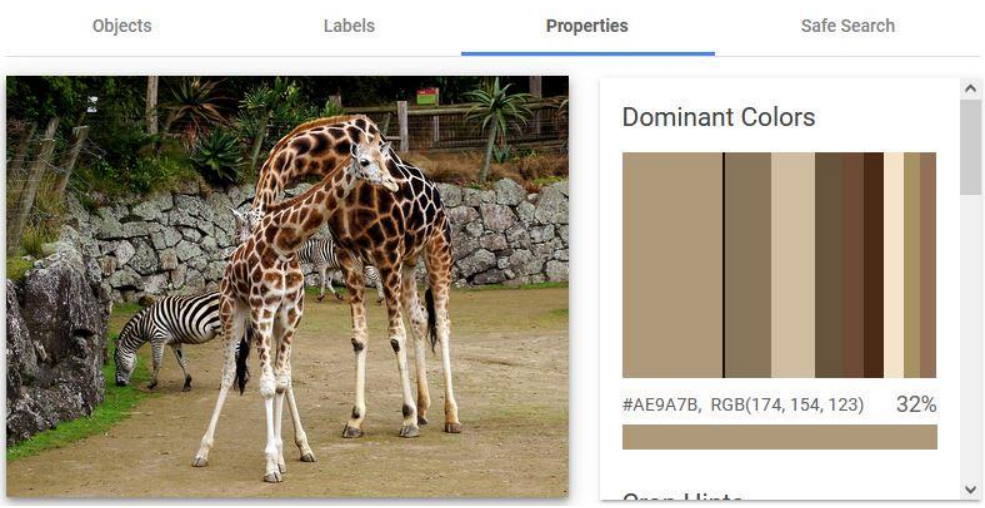
```
    "topicality": 0.5251859
  },
  {
    "description": "Safari",
    "mid": "/m/0199b0",
    "score": 0.51667583,
    "topicality": 0.51667583
  },
  {
    "description": "Pasture",
    "mid": "/m/01ɡd91",
    "score": 0.51553565,
    "topicality": 0.51553565
  },
  {
    "description": "Savanna",
    "mid": "/m/01c7d3",
    "score": 0.5093298,
    "topicality": 0.5093298
  },
  {
    "description": "Natural landscape",
    "mid": "/m/03d28y3",
    "score": 0.50280577,
    "topicality": 0.50280577
  },
  {
    "description": "Herd",
    "mid": "/m/03vd46",
    "score": 0.5009309,
    "topicality": 0.5009309
  }
],
```



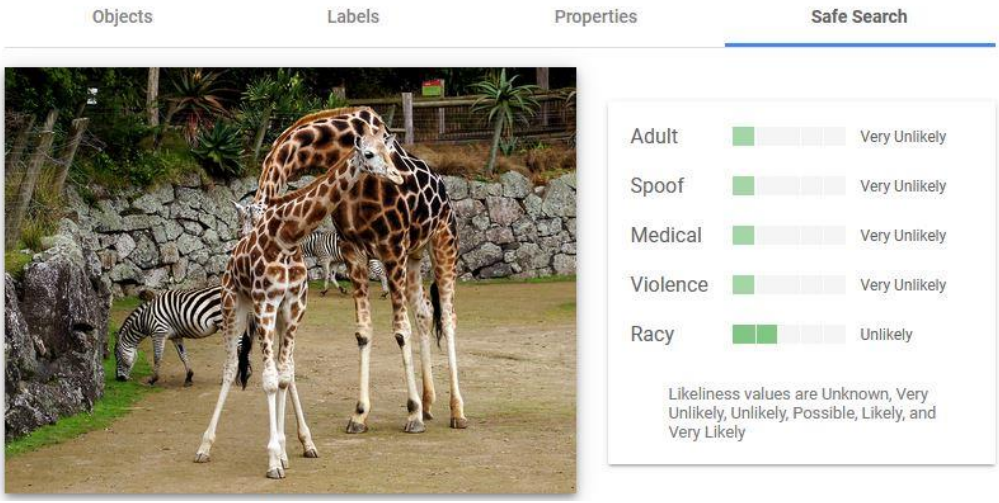
ภาพที่ 27 ตัวอย่างผลการตรวจจับวัตถุในภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ



ภาพที่ 28 ตัวอย่างผลการรู้จำภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ

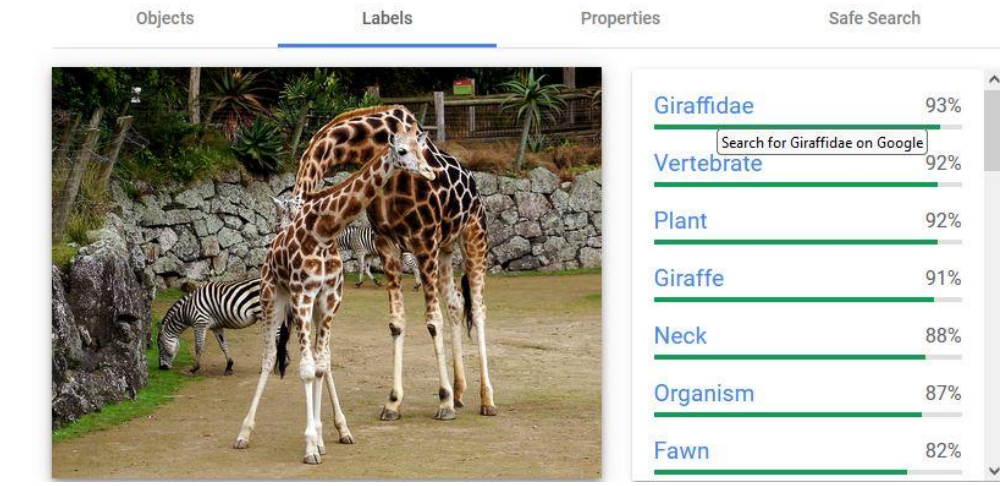


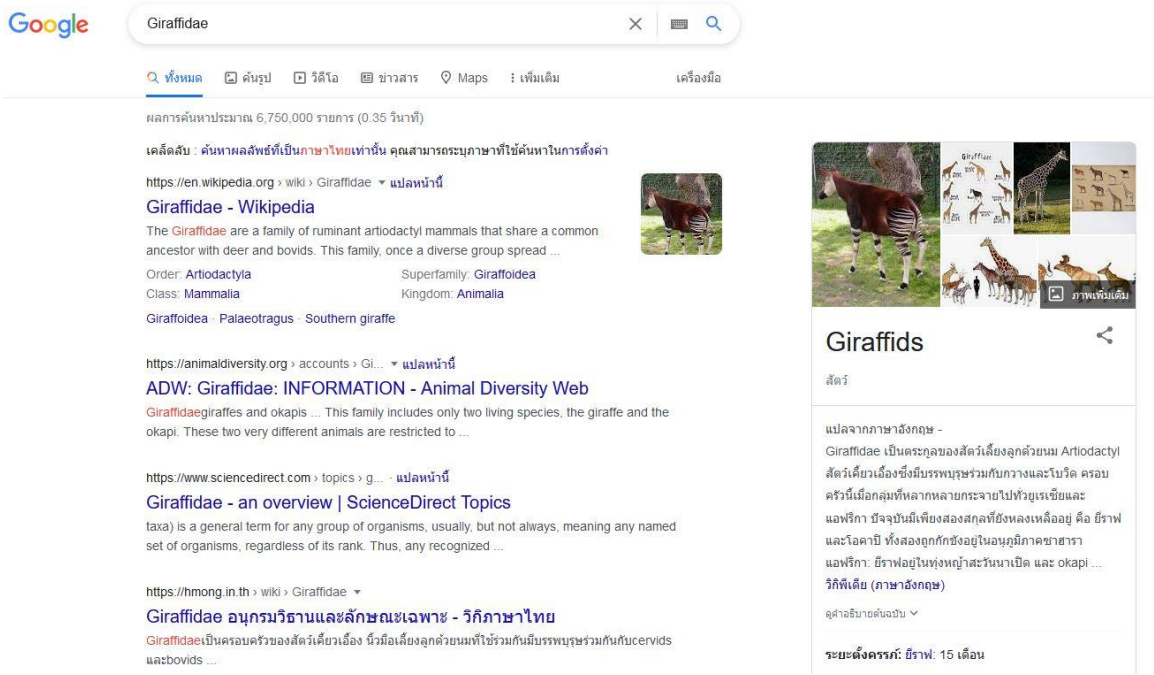
ภาพที่ 29 ตัวอย่างผลการรู้จำสีในภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ



ภาพที่ 30 ตัวอย่างผลการรู้จำความปลอดภัยในภาพถ่ายด้วยโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ

สามารถคลิกที่ผลลัพธ์ที่ได้จากการรู้จำในภาพเพื่อเชื่อมโยงการค้นหาที่เกี่ยวข้องในกูเกิลได้ ทำให้สามารถค้นหาความหมายเพิ่มเติมคำศัพท์เมทาดาทาได้อย่างรวดเร็ว





ภาพที่ 31 การเชื่อมโยงผลลัพธ์จากโปรแกรมกูเกิลคลาวน์วิชันเอพีไอไปยังโปรแกรมกูเกิล

บันทึกเมทาดาทาที่ได้จากการใช้เครื่องมือ

แบบบันทึกข้อมูลเมทาดาทาจากเครื่องมือรู้จำภาพด้วยเทคนิคปัญญาประดิษฐ์ ประกอบด้วย รหัสไฟล์ภาพ (File ID) วันที่ทดสอบ (Test date) ข้อมูลเมทาดาทา (Metadata) พร้อมค่าระดับความเชื่อมั่น (Confidence score) จากโปรแกรมอเมซอน เรคคอคนิชัน คลาริฟาย อิมเมกกา และกูเกิลคลาวน์วิชันเอพีไอ โดยใช้เมทาดาทาที่มีค่าระดับความเชื่อมั่นมากกว่า 50% หรือในกรณีที่จำนวนเมทาดาทาที่มีค่าระดับความเชื่อมั่นมากกว่า 50% จำนวนน้อย จะใช้เมทาดาทา 10 อันดับแรกที่ได้ ดังแสดงในตารางที่ 4 ถึง ตารางที่ 7

การเปรียบเทียบเมทาดาทาเรียงลำดับตามค่าความเชื่อมั่นแสดงในตารางที่ 8 และการเปรียบเทียบเมทาดาทาเรียงลำดับตามตัวอักษรแสดงในตารางที่ 9

ตารางที่ 4 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมอเมซอน เรคคอคนิชันของภาพรหัส A010

รหัสไฟล์ภาพ (File ID)	**วันที่ทดสอบ (Test date)**	**ข้อมูลเมทาดาทา (Metadata)**	**ค่าระดับความเชื่อมั่น (Confidence score: %)**
A010	สิงหาคม 2565	zebra	98.10
		mammal	98.10
		animal	98.10
		wildlife	98.10
		giraffe	97.30
		zoo	61.70

ตารางที่ 5 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมอิมเมกกาของภาพรหัส A010

รหัสไฟล์ภาพ (File ID)	**วันที่ทดสอบ (Test date)**	**ข้อมูลเมทาดาทา (Metadata)**	**ค่าระดับความเชื่อมั่น (Confidence score: %)**
A010	สิงหาคม 2565	giraffe	100.00
		wildlife	68.80
		safari	64.63
		animal	60.09
		mammal	57.66
		wild	56.74
		zebra	54.99
		zoo	41.61
		park	32.19
		reserve	29.17

ตารางที่ 6 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมคลาริฟายของภาพรหัส A010

รหัสไฟล์ภาพ (File ID)	**วันที่ทดสอบ (Test date)**	**ข้อมูลเมทาดาทา (Metadata)**	**ค่าระดับความเชื่อมั่น (Confidence score: %)**
A010	สิงหาคม 2565	zoo	99.20
		safari	99.20
		animal	99.10
		wildlife	99.00
		giraffe	98.50
		nature	98.20
		mammal	98.00
		wild	97.60
		grass	96.70
		neck	95.20
		savanna	91.70
		park	91.60
		reserve	90.00
		herbivore	88.30
		barbaric	87.40
		no person	85.00
		jungle	84.80
		long	84.60
		tree	83.20
		head	82.90

ตารางที่ 7 แบบบันทึกข้อมูลเมทาดาทาจากโปรแกรมกูเกิลคลาวน์วิชันเอพีไอของภาพรหัส A010

รหัสไฟล์ภาพ (File ID)	**วันที่ทดสอบ (Test date)**	**ข้อมูลเมทาดาทา (Metadata)**	**ค่าระดับความเชื่อมั่น (Confidence score: %)**
A010	สิงหาคม 2565	giraffidae	93.74
		giraffe	93.61
		plant	93.44
		vertebrate	92.00
		neck	88.16
		organism	86.84
		zebra	82.85
		fawn	81.60
		grass	81.56
		terrestrial animal	81.55
		landscape	79.40
		adaptation	79.20
		recreation	77.51
		snout	76.98
		grassland	76.88
		wildlife	71.52
		pattern	67.15
		nature reserve	66.62
		tree	62.17
		jungle	61.55
		tail	57.91
		palm tree	57.58
		zoo	55.05
		groundcover	53.95
		pasture	53.75
		savanna	52.85
		safari	52.61
		trunk	52.44
		natural landscape	52.06
		herd	51.61

ตารางที่ 8 เปรียบเทียบเมทาดาทาเรียงลำดับตามค่าความเชื่อมั่น

เครื่องมือ	**จำนวนคำ**	**เมทาดาทาเรียงลำดับตามค่าความเชื่อมั่น**
มาพร้อมภาพถ่าย	5	animal, creatures, giraffe, zebra, zoo
อเมซอน เรคคอคนิชัน	6	zebra, mammal, animal, wildlife, giraffe, zoo
คลาริฟาย	20	zoo, safari, animal, wildlife, giraffe, nature, mammal, wild, grass, neck, savanna, park, reserve, herbivore, barbaric, no person, jungle, long, tree, head
อิมเมกกา	10	giraffe, wildlife , safari, animal, mammal, wild, zebra, zoo, park, reserve
กูเกิลคลาวน์วิชันเอพีไอ	31	giraffidae , giraffe, plant, vertebrate,neck, organism, zebra, fawn, grass, terrestrial, animal, landscape, adaptation, recreation, snout, grassland, wildlife pattern, nature reserve, tree, jungle, tail, palm tree , zoo, groundcover, pasture, savanna, safari, trunk, natural, landscape, herd

ตารางที่ 9 เปรียบเทียบเมทาดาทาเรียงลำดับตามตัวอักษร

เครื่องมือ	**เมทาดาทาเรียงลำดับตามตัวอักษร**
มาพร้อมภาพถ่าย	animal, creatures, giraffe, zebra, zoo
อเมซอน เรคคอคนิชัน	**animal, giraffe,** mammal, wildlife, **zebra, zoo**
คลาริฟาย	**animal,** barbaric, **giraffe,** grass, herbivore, jungle, long, mammal, nature, neck, no person, park, reserve, safari, savanna, tree, wild, wildlife, **zoo**
อิมเมกกา	**animal, giraffe,** mammal, park, reserve, safari, wild, wildlife, **zebra, zoo**
กูเกิลคลาวน์วิชันเอพีไอ	adaptation, animal, fawn, **giraffe,** giraffidae, grass, grassland, groundcover, herd jungle, landscape, natural, nature reserve, neck, organism, palm tree, pasture, pattern, plant, recreation, safari, savanna safari, snout, tail, terrestrial, **animal**, tree trunk, vertebrate, wildlife, **zebra, zoo**

การวิเคราะห์ประสิทธิภาพ

การวิเคราะห์ประสิทธิภาพของเครื่องมือการรู้จำภาพถ่ายนี้ใช้วิธีการทั้งเชิงปริมาณและคุณภาพ ในเชิงปริมาณจะตรวจสอบจำนวนภาพถ่ายที่สามารถแยกเมทาดาทาได้ค่าความถูกต้องมากกว่า 80% ในเชิงคุณภาพใช้วิธีการวิเคราะห์ข้อมูลสองแบบ คือ วิธีการทางสถิติเพื่อหาค่าความคล้ายคลึงของเอกสาร (Similarity) โดยใช้การคำนวณค่าความคล้ายโคไซน์ (Cosine Similarity) ค่าคำทับซ้อน (Word Overlap Similarity Measure) และวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap SimilarityMeasure) ส่วนการหาค่าประสิทธิภาพของเครื่องมือ ใช้การคำนวณหาค่าความไว (Recall) ค่าความแม่นยำ (Precision) และ ค่าความถูกต้องเอฟวัน (F1-score) (Hancock, 2004; Singthongchai & Niwattanakul, 2013; Abdullah el al., 2020)

การวัดค่าความถูกต้องของเมทาดาทาทำโดยการนำเมทาดาทาที่ได้มาพร้อมกับภาพถ่ายเปรียบเทียบกับเมทาดาทาที่ได้จากโปรแกรม เพื่อหาค่าความคล้ายคลึง หากโปรแกรมใดมีคำที่มีความคล้ายคลึงเข้าใกล้ 1 มากที่สุดหมายความว่ามีเมทาดาทาที่เหมือนกันมากที่สุด และวัดค่าประสิทธิภาพของเครื่องมือ จากการค่าการหาค่าความไว ค่าความแม่นยำ และ ค่าความถูกต้องเอฟวัน

การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)

การวัดค่าความทับซ้อนเป็นการคำนวณจำนวนคำที่มีร่วมกัน คำนวณได้จากจำนวนคำเมทาดาทาที่มีร่วมกันจากแต่ละเครื่องมือ หารด้วยจำนวนคำเมทาดาทาที่ได้มาพร้อมภาพถ่าย ดังสมการที่ (1)

$$sim_{overlap}(A, B) = \frac{|A \cap B|}{|A|} \quad (1)$$

โดย $|A|$ หมายถึง จำนวนคำเมทาดาทาที่ได้มาพร้อมภาพถ่าย

$|B|$ หมายถึง จำนวนคำเมทาดาทาที่ได้มาจากเครื่องมือ

$|A \cap B|$ หมายถึง จำนวนคำเมทาดาทาที่เครื่องมือมีร่วมกัน

ตารางที่ 10 ผลการวัดค่าคำที่ทับซ้อนของภาพรหัส A010

เครื่องมือ	จำนวนคำ	เมทาดาทาที่ทับซ้อนกัน
มาพร้อมภาพถ่าย	5	animal, creatures, giraffe, zebra, zoo
อเมซอน เรคคอคนิชัน	4	animal, giraffe, zebra, zoo
คลาริฟาย	4	animal, giraffe, zebra, zoo
อิมเมกกา	4	animal, giraffe, zebra, zoo
กูเกิลคลาวน์วิชันเอพีไอ	4	animal, giraffe, zebra, zoo

ในกรณีของตัวอย่างภาพ A010 เมทาดาทาที่มาพร้อมภาพถ่ายมีจำนวน 5 คำ ทุกโปรแกรมมีคำทับซ้อนเท่ากับคือ 4 คำ จึงมีค่าคำทับซ้อนเท่ากับ $\frac{4}{5}$ = 0.80 หรือ 80% เท่ากันทุกโปรแกรม

การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap SimilarityMeasure)

เป็นการวัดค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม หากนำมาใช้ในการกำหนดเมทาดาทาแล้วจะสามารถเพิ่มโอกาสในการค้นหาภาพได้มากยิ่งขึ้น คำนวณได้จากจำนวนคำเมทาดาทาที่มีไม่มีร่วมกันแต่ให้ความหมายที่เกี่ยวข้องกับภาพจากแต่ละเครื่องมือ หารด้วยจำนวนคำเมทาดาทาที่ได้มาพร้อมภาพถ่าย ค่าที่ได้จะแสดงโอกาสในการค้นหาภาพที่มากยิ่งขึ้น ดังสมการที่ (2)

$$sim_{NoOverlap}(A,B) = \frac{|B_{NoOverlap}|}{|A|} \quad (2)$$

โดย $|B_{NoOverlap}|$ หมายถึงจำนวนคำคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม

$|A|$ หมายถึง จำนวนคำเมทาดาทาที่ได้มาพร้อมภาพถ่าย

ตารางที่ 11 ผลการวัดค่าคำที่ไม่ทับซ้อนของภาพรหัส A010

เครื่องมือ	**จำนวนคำ**	**เมทาดาทาที่มีความหมายเกี่ยวข้อง**
มาพร้อมภาพถ่าย	5	animal, creatures, giraffe, zebra, zoo
อเมซอน เรคคอคนิชัน	2	mammal, wildlife
คลาริฟาย	4	mammal, wildlife, grass, safari
อิมเมกกา	3	mammal, wildlife, safari
กูเกิลคลาวน์วิชันเอพีไอ	8	wildlife, vertebrate, safari, grass, plant, recreation, animal, palm tree, giraffidae

ค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม ของแต่ละโปรแกรมมีค่าดังนี้

อเมซอน เรคคอคนิชัน = $\frac{2}{5}$ = 0.40 หรือ โอกาสในการค้นหาเพิ่มขึ้น 40%

คลาริฟาย = $\frac{4}{5}$ = 0.80 หรือ โอกาสในการค้นหาเพิ่มขึ้น 80%

อิมเมกกา = $\frac{3}{5}$ = 0.60 หรือ โอกาสในการค้นหาเพิ่มขึ้น 60%

กูเกิลคลาวน์วิชันเอพีไอ = $\frac{8}{5}$ = 1.60 หรือ โอกาสในการค้นหาเพิ่มขึ้น 160%

การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)

การหาค่าความคล้ายคลึงใช้วิธีหาค่าสัมประสิทธิ์ความละม้ายโคไซน์เป็นการหาค่าความคล้ายของเมทาดาทาได้กำกับมาพร้อมภาพถ่ายว่าเหมือนกันกับเมทาดาทาที่ได้จากเครื่องมือมากน้อยแค่ไหน เหมาะกับการหาค่าความคล้ายของคำที่มีขนาดไม่เท่ากัน ค่าที่ได้จากเป็นตัวเลขตั้งแต่ 0 ถึง 1 หากมีค่าเท่ากับ 0 แปลว่าไม่มีความคล้ายกันเลย ถ้ามีค่าความคล้ายเข้าใกล้ 1 แปลว่ามีความคล้ายกันสูง คำนวณได้จากการนาผลคูณ (dot product) ของเวกเตอร์ 2 เวกเตอร์แล้วนำมาหารด้วยขนาดของทั้งสองเวกเตอร์คำนวณได้จากสมการ (3) ดังนี้

$$cos(\emptyset) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i\, B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \quad (3)$$

โดย *A* และ *B* คือเวกเตอร์ที่แสดงเมทาดาทาของสองเครื่องมือ

$A \cdot B$ คือ ผลรวมของผลคูณขององค์ประกอบที่สอดคล้องกันของเวกเตอร์

$\|A\|$ และ $\|B\|$ คือค่าแมกนิจูดของเวกเตอร์ของเมทาดาทาแต่ละเครื่องมือ

ตารางที่ 12 ผลการวัดค่าความคล้ายโคไซน์ ของภาพรหัส A010

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
มาพร้อมภาพถ่าย	5	animal, creatures, giraffe, zebra, zoo
อเมซอน เรคคอคนิชัน	6	animal, giraffe, mammal, wildlife, zebra, zoo
คลาริฟาย	20	animal, barbaric, giraffe, grass, herbivore, jungle, long, mammal, nature, neck, no person, park, reserve, safari, savanna, tree, wild, wildlife, zoo
อิมเมกกา	10	animal, giraffe, mammal, park, reserve, safari, wild, wildlife, zebra, zoo
กูเกิลคลาวน์วิชันเอพีไอ	31	adaptation, animal, fawn, giraffe, giraffidae, grass, grassland, groundcover, herd, jungle, landscape, natural, nature reserve, neck, organism, palm tree, pasture, pattern, plant, recreation, safari, savanna, safari, snout, tail, terrestrial, animal, tree, trunk, vertebrate, wildlife, zebra, zoo

ค่าสัมประสิทธิ์ความคล้ายโคซายน์ของแต่ละโปรแกรมมีค่าดังนี้

อเมซอน เรคคอคนิชัน $= \frac{4}{\sqrt{5\sqrt{6}}} = 0.73$ หรือ 73%

คลาริฟาย $= \frac{4}{\sqrt{5\sqrt{20}}} = 0.40$ หรือ 40%

อิมเมกกา $= \frac{4}{\sqrt{5\sqrt{10}}} = 0.57$ หรือ 57%

กูเกิลคลาวน์วิชันเอพีไอ $= \frac{4}{\sqrt{5\sqrt{31}}} = 0.32$ หรือ 32%

วัดค่าประสิทธิภาพของเครื่องมือ

การวัดค่าประสิทธิภาพของเมทาดาทาที่ได้จากเครื่องมือใช้เกณฑ์วัดประสิทธิภาพของเครื่องมือ 3 ค่า ได้แก่ ค่าความไว (Recall) ค่าความแม่นยำ และ ค่าความถูกต้องเอฟวัน (F1-score) เป็นการคำนวณประสิทธิภาพ

ค่าความไว (Recall) คือค่าจำนวนคำเมทาดาทาที่เครื่องมือกำหนดได้ถูกต้อง และถูกต้องตามค่าจริง (True Positive) ของข้อมูลต่อจำนวนเมทาดาทาทั้งหมดที่ถูกต้อง (True Positive + False Negative)

ค่าความแม่นยำ (Precision) คือค่าจำนวนคำเมทาดาทาที่เครื่องมือกำหนดได้ถูกต้อง และถูกต้องตามค่าจริง (True Positive) ของข้อมูลต่อจำนวนเมทาดาทาที่เครื่องมือกำหนดได้ทั้งหมด (True Positive + False Positive)

ค่าความถูกต้องเอฟวัน (F1-score) คือตัววัดค่าความสัมพันธ์ หรือค่าเฉลี่ยของค่าความไวและค่าความแม่นยำ ใช้เป็นตัววัดประสิทธิภาพของเครื่องมือที่กำหนดเมทาดาทา ซึ่งมีค่าอยู่ระหว่าง 0 ถึง 1 โดยค่า F1-score เข้าใกล้ 1 หมายถึงมีเครื่องมือมีประสิทธิภาพสูง คำนวณจากสมการ (4)-(6) ดังนี้

โดยทั่วไปเมื่อประเมินเครื่องมือสร้างเมทาดาทา ควรพิจารณาทั้งค่าความไว และค่าความแม่นยำด้วยเพื่อให้ได้ความเข้าใจทั้งหมดเกี่ยวกับประสิทธิภาพของเครื่องมือ เครื่องมือที่มีค่าความแม่นยำสูงแต่ค่าความไวต่ำอาจจะดีที่สามารถระบุเมทาดาทาได้ถูกต้องจากที่สร้างทั้งหมด แต่อาจไม่พบเมทาดาทาที่ถูกต้องตามค่าจริงทั้งหมด ส่วนเครื่องมือที่มีค่าความไวสูงแต่ ค่าความแม่นยำต่ำคือสามารถกำหนดเมทาดาทาได้ถูกต้องจำนวนมากจากเมทาดาทาตามค่าจริง อีกอย่างหนึ่งคือเครื่องมือที่มีความไวสูงแต่ความแม่นยำต่ำอาจจะระบุเมทาดาทาที่ใกล้เคียงแต่อาจจะเกิดจำนวนของเมทาดาทาที่สร้างได้แต่ไม่ถูกต้อง (false positive) สูง ดังนั้นค่าความถูกต้องเอฟวัน จึงเป็นเกณฑ์ที่ดีในการเปรียบเทียบความสมดุลระหว่างความแม่นยำและความไว

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

$$\text{F1-score} = 2 * \frac{precision * recall}{precision + recall} \quad (6)$$

โดย TP หมายถึงเมทาดาทาที่พบถูกต้อง FN หมายถึงเมทาดาทาที่พบแต่ไม่ถูกต้อง FP หมายถึงเมทาดาทาที่ถูกต้องแต่ไม่พบ

ตารางที่ 13 ผลการวัดค่าความถูกต้องเมทาดาทาของภาพรหัส A010

เครื่องมือ	**TP**	**FN**	**FP**	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	4	2	1	0.67	0.80	0.73
คลาริฟาย	4	16	1	0.20	0.80	0.32
อิมเมกกา	4	6	1	0.40	0.80	0.53
กูเกิลคลาวน์วิชันเอพีไอ	4	27	1	0.13	0.80	0.22

จากวิธีการวัดค่าต้นจะเห็นว่าเป็นการวัดค่าความละม้ายของเมทาดาทาที่มีคำตรงกันเท่านั้น คำที่มีความหมายใกล้เคียงไม่ได้นำมาพิจารณาร่วมด้วย ค่าความละม้ายที่ได้ออกมาจึงไม่ตรงกับความเป็นจริงเท่าใดนัก เพื่อให้งานวิจัยนี้มีความน่าเชื่อถือเพิ่มขึ้น ควรพิจารณาความคล้ายคลึงในเชิงความหมาย (semantic similarity) ร่วมด้วย

บทที่ 4

ผลการทดลอง

วัตถุประสงค์งานวิจัยนี้เพื่อสำรวจเครื่องมือในการรู้จำวัตถุในภาพถ่ายที่ใช้ปัญญาประดิษฐ์เพื่อกำหนดเมทาดาท และ เพื่อวิเคราะห์และประเมินประสิทธิภาพเครื่องมือในการรู้จำภาพที่เหมาะสมกับภาพถ่ายแต่ละประเภท จากการดำเนินการวิจัยในบทที่ 3 ซึ่งแบ่งเป็น 5 ขั้นตอนหลัก คือ การเก็บรวบรวมภาพถ่าย การบันทึกเมทาดาทาที่มาพร้อมภาพถ่าย การทดสอบเครื่องมือการรู้จำภาพถ่ายด้วยเทคนิคปัญญาประดิษฐ์ การบันทึกเมทาดาทาที่ได้จากเครื่องมือปัญญาประดิษฐ์ และการวิเคราะห์ประสิทธิภาพเมทาดาทาและเครื่องมือปัญญาประดิษฐ์

ในการวิเคราะห์และประเมินประสิทธิภาพเครื่องมือ ใช้วิธีการทางสถิติเพื่อหาค่าความคล้ายคลึงของเอกสาร (Similarity) ซึ่งประกอบด้วย การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure) การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap SimilarityMeasure) การวัดค่าความคล้ายโคไซน์ (Cosine Similarity) วัดค่าประสิทธิภาพของเครื่องมือ การหาค่าความไว (Recall) ค่าความแม่นยำ (Precision) และ ค่าความถูกต้องเอฟวัน (F1-score) ผลการทดลองแสดงได้ดังต่อไปนี้

ผลการเก็บรวบรวมภาพถ่าย พร้อมบันทึกเมทาดาทาที่มาพร้อมภาพถ่าย

เก็บรวบรวมภาพถ่ายจากคลังภาพดิจิทัลแบบเปิด เผยแพร่ภายใต้โดเมนสาธารณะ ซึ่งไม่มีลิขสิทธิ์หรือเจ้าของลิขสิทธิ์ยกให้เป็นสาธารณะประโยชน์ จำนวน 1,000 ภาพ โดยแบ่งหมวดหมู่ภาพถ่ายเป็น 10 หมวดหมู่ๆ ละ 100 ภาพ ประกอบด้วย

1. ภาพคน (human) ให้รหัสภาพถ่ายขึ้นต้นด้วย H
2. ภาพสัตว์ (animal) ให้รหัสภาพถ่ายขึ้นต้นด้วย A
3. ภาพต้นไม้ ดอกไม้ (plant and flower) ให้รหัสภาพถ่ายขึ้นต้นด้วย PF
4. ภาพวิวทิวทัศน์ (view and landscape) ให้รหัสภาพถ่ายขึ้นต้นด้วย VL
5. ภาพผัก ผลไม้ (vegetable and fruit) ให้รหัสภาพถ่ายขึ้นต้นด้วย VF
6. ภาพอาหาร (food) ให้รหัสภาพถ่ายขึ้นต้นด้วย F
7. ภาพยานพาหนะ (vehicle) ให้รหัสภาพถ่ายขึ้นต้นด้วย V
8. ภาพสถานที่ท่องเที่ยว (tourist landmark) ให้รหัสภาพถ่ายขึ้นต้นด้วย L
9. ภาพศิลปวัฒนธรรม (art and culture) ให้รหัสภาพถ่ายขึ้นต้นด้วย AC
10. ภาพปกหนังสือ/โปสเตอร์เก่า (old book cover and poster) ให้รหัสภาพถ่ายขึ้นต้นด้วย BP

ภาพถ่ายที่ใช้ในการทดลองทั้งหมด บันทึกเมทาดาทาที่ได้มาพร้อมภาพถ่าย และบันทึกเมทาดาทาที่ได้จากเครื่องมือ สามารถเข้าถึงได้ที่ website https://sites.google.com/view/ajakara/research

ผลการทดสอบเครื่องมือปัญญาประดิษฐ์ และบันทึกเมทาดาทาที่ได้จากเครื่องมือ

ทำการทดลองกับภาพถ่ายทั้งหมดจำนวน 1,000 ภาพแบ่งตามหมวดหมู่ที่กำหนด แสดงเมทาดาทาที่ได้มาพร้อมภาพถ่ายเปรียบเทียบกับเมทาดาทาที่ได้จากเครื่องมือปัญญาประดิษฐ์ 4 โปรแกรม คือ อเมซอนเรคคอคนิชัน คลาริฟาย อิมเมกกา และกูเกิลคลาวน์วิชันเอพีไอ โดยยกตัวอย่างผลการทดลองของภาพถ่ายมาแสดงหมวดหมู่ละ 3 ภาพ

หมวดหมู่ภาพคน

ตารางที่ 14 เมทาดาทาจากการทดลองในหมวดหมู่ภาพคน รหัสภาพ H01

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ H01	7	action, alpine, Korea, skiing, snow, sport, winter
อเมซอน เรคคอคนิชัน	20	spire, steeple, building, tower, architecture, metropolis, city, urban, town, landscape, nature, outdoors, scenery, neighborhood, aerial view, downtown, azure sky, sky, high rise
คลาริฟาย	20	snow, winter, skier, competition, race, goggles, ice, action, downhill, action energy, fast, gloves, sport, cold, resort, helmet, man, hurry, alpine, fun
อิมเมกกา	10	ski, snow, winter, skier, cold, sport, fun, mountain, outdoor, season
กูเกิลคลาวน์วิชันเอพีไอ	26	footwear, trousers, helmet, sports equipment, shoe, glove, sports uniform, snow, white, sports gear, ski equipment, goggles, ski, winter sport, outdoor recreation, slope, headgear, ski pole, ski boot, skier, eyewear, sneakers, competition event, freezing, personal protective equipment, cross-country skiing, sports, ski helmet, recreation, hat, glacial landform, competition, slalom skiing, ice cap, individual sports, ski cross, fun, sportswear, winter, skiing, piste, leisure, ski binding, exercise, jumping, nordic skiing, ski touring, playing in the snow, nordic combined

ตารางที่ 15 เมทาดาทาจากการทดลองในหมวดหมู่ภาพคน รหัสภาพ H02

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ H02	7	animal, dog, golden retriever, friend, love, woman
อเมซอน เรคคอคนิชัน	22	person, human, dog, canine, animal, pet, mammal, couch, furniture, doctor, golden retriever, home decor, monitor, electronics, screen, display, lcd screen, people, female
คลาริฟาย	20	dog, pet, canine, people, woman, mammal, owner, animal, girl, portrait, friendship, cute, family, puppy, adult, love, domestic, sit, blond, elder
อิมเมกกา	12	corgi, dog, canine, domestic animal, dingo, pet, wild dog, cute, fur, cat
กูเกิลคลาวน์วิชันเอพีไอ	34	dog, carnivore, comfort, companion dog, television, service, flooring, dog breed, home appliance, veterinarian, working animal, event, sporting group, wood, fur, television set, leisure, room, cabinetry, entertainment center, sitting, fence, machine, led-backlit lcd display, canidae

ตารางที่ 16 เมทาดาทาจากการทดลองในหมวดหมู่ภาพคน รหัสภาพ H11

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ H11	9	casual, computer, couch, cropped, girl, laptop, man, seat, sit
อเมซอน เรคคอคนิชัน	20	couch, furniture, person, human, laptop, computer, pc, electronics, cushion, dating, pants, clothing, apparel, sitting, heel, female, girl, video gaming, people
คลาริฟาย	20	sofa, woman, sit, family, seat, indoors, girl, man, people, room, internet, wireless, casual, relaxation, furniture, adult, leisure, computer, college, technology
อิมเมกกา	10	miniskirt, adult, laptop, person, sofa, sitting, home, skirt, computer, couch
กูเกิลคลาวน์วิชันเอพีไอ	49	computer, white, laptop, leg, personal computer, comfort, sleeve, gadget, thigh, knee, travel, shorts, t-shirt, leisure, eyewear, lap, electronic device, human leg, white-collar worker, netbook, electric blue, fun, sitting, job, plaid, uniform, bag, room, conversation, reading, pattern, sportswear, employment, vacation, public transport, passenger, learning, luggage and bags, tourism, luxury vehicle

หมวดหมู่ภาพสัตว์

ตารางที่ 17 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสัตว์ รหัสภาพ A04

เครื่องมือ	จำนวนคำ	เมทาดาทา
 รหัสภาพ A04	6	animal, zoo, monkey, gorilla, mother, baby
อเมซอน เรคคอคนิชัน	5	gorilla, animal, mammal, wildlife, ape
คลาริฟาย	21	ape, primate, monkey, gorilla, mammal, wildlife, zoo, animal, captive, chimpanzee, endangered, portrait, hairy, jungle, endangered species, baby, strong, fur, safari, sumatra
อิมเมกกา	10	gorilla, ape, primate, monkey, wildlife, wild, zoo, animals, endangered, natural
กูเกิลคลาวน์วิชันเอพีไอ	23	head, primate, eye, grass, terrestrial animal, common chimpanzee, snout, thumb, natural material, grassland, fur, wildlife, wrinkle, sitting, jungle, happy, groundcover, zoo, comfort, paw

ตารางที่ 18 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสัตว์ รหัสภาพ A06

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ A06	6	animal, zoo, polar bear, swim, fish, eating
อเมซอน เรคคอคนิชัน	6	bear, mammal, wildlife, animal, polar bear
คลาริฟาย	21	water, swimming, zoo, wildlife, sea, mammal, ocean, underwater, animal, fish, nature, marine, dug-out pool, influencer, aquarium, portrait, wild, frosty, river, lake
อิมเมกกา	11	ice bear, bear, mammal, wildlife, water, wild, fur, polar, zoo, sea
กูเกิลคลาวน์วิชันเอพีไอ	18	water, polar bear, liquid, vertebrate, carnivore, fluid, organism, mammal, drink, terrestrial animal, snout, ear, marine mammal, wildlife, ocean

ตารางที่ 19 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสัตว์ รหัสภาพ A11

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ A11	5	giant, turtle, animal, reptile, tortoise
อเมซอน เรคคอคนิชัน	8	turtle, reptile, sea life, animal, tortoise, box turtle
คลาริฟาย	21	reptile, wildlife, turtle, nature, animal, tortoise, slow, zoo, no person, shell, wild, portrait, water, eye, amphibian, tropical, skin, scale, exotic, zoology
อิมเมกกา	11	turtle, reptile, mud turtle, terrapin, slow, shell, tortoise, wildlife, wild, pet
กูเกิลคลาวน์วิชันเอพีไอ	31	head, eye, reptile, human body, chelonoidis, gopher tortoise, turtle, galápagos tortoise, desert tortoise, terrestrial animal, tortoise, snout, pond turtle, grass, electric blue, wildlife, macro photography, gopher tortoise, scaled reptile, biting, geoemydidae

หมวดหมู่ภาพต้นไม้ ดอกไม้

ตารางที่ 20 เมทาดาทาจากการทดลองในหมวดหมู่ภาพต้นไม้ ดอกไม้ รหัสภาพ PF01

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ PF01	6	bloom, flower, garden, love, pink, rose
อเมซอน เรคคอคนิชัน	5	rose, plant, flower, blossom, petal
คลาริฟาย	19	nature, leaf, romance, flora, love, petal, floral, blooming, no person, garden, bouquet, bud, romantic, bright, outdoors, gift, summer, shrub
อิมเมกกา	9	shrub, flower, plant, pink, petal, roses, blossom, love, floral
กูเกิลคลาวน์วิชันเอพีไอ	50	botany, leaf, nature, natural environment, hybrid tea rose, petal, branch, pink, grass, rose, garden roses, rosa × centifolia, beauty, flowering plant, shrub, annual plant, rose family, close-up, rose order, floribunda, peach, herbaceous plant, spring, groundcover, magenta, china rose, plant stem, perennial plant, garden, begonia family, carmine, landscape, rosa wichuraiana, begonia

ตารางที่ 21 เมทาดาทาจากการทดลองในหมวดหมู่ภาพต้นไม้ ดอกไม้ รหัสภาพ PF02

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ PF02	6	flower, floral, sky, sun, sunflower, yellow
อเมซอน เรคคอคนิชัน	4	plant, flower, blossom, sunflower
คลาริฟาย	19	nature, flower, summer, flora, pollen, bright, growth, leaf, petal, fair weather, stamen, outdoors, field, garden, vibrant, floral, sun, rural
อิมเมกกา	9	sunflower, flower, yellow, plant, field, summer, petal, sun, agriculture, flora
กูเกิลคลาวน์วิชันเอพีไอ	29	sky, petal, yellow, cuisine, flowering plant, annual plant, sunflower, close-up, pollen, field, herbaceous plant, plantation, event, wildflower, forb, macro photography, sunflower, asterales, daisy family, vegetarian food, pedicel, plant stem

ตารางที่ 22 เมทาดาทาจากการทดลองในหมวดหมู่ภาพต้นไม้ ดอกไม้ รหัสภาพ PF04

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ PF04	6	flower, leaf, lotus, magenta, pink, pond
อเมซอน เรคคอคนิชัน	5	lily, plant, flower, blossom, pond lily
คลาริฟาย	19	nature, lily, flower, exotic, tropical, aquatic, leaf, no person, waterlily, water, meditation, summer, sacred, zen, flora, nelumbo, lake, outdoors
อิมเมกกา	9	flower, pink, blossom, plant, petal, flora, bloom, garden, clover
กูเกิลคลาวน์วิชันเอพีไอ	51	plant, lotus, natural environment, botany, organism, sacred lotus, vegetation, terrestrial plant, petal, biome, pink, lake, natural landscape, groundcover, aquatic plant, flowering plant, fragrant white water lily, annual plant, lotus family, landscape, magenta, water lily, reflection, wetland, wildflower, proteales, plant stem, liquid, macro photography, dicotyledon, pond, wildlife, herbaceous plant, lily family, garden

หมวดหมู่ภาพวิวทิวทัศน์

ตารางที่ 23 เมทาดาทาจากการทดลองในหมวดหมู่ภาพวิวทิวทัศน์ รหัสภาพ VL05

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ VL05	5	fuji mountain, japan, landscape, mountain, nature
อเมซอน เรคคอคนิชัน	14	nature, mountain range, outdoors, mountain, scenery, peak, plateau, landscape, land, wilderness, tree, plant, panoramic
คลาริฟาย	21	volcano, snow, mountain, sunset, dawn, landscape, travel, sky, no person, nature, winter, sun, fog, lake, evening, outdoors, dusk, light, mist, fall
อิมเมกกา	12	volcano, mountain, natural elevation, geological formation, landscape, mountains, peak, snow, sky, travel
กูเกิลคลาวน์วิชันเอพีไอ	65	sky, mountain, cloud, nature, highland, snow, water, natural landscape, tree, body of water, slope, sunlight, atmospheric phenomenon, plant, horizon, landscape, mountainous landforms, morning, mountain range, hill, volcano, calm, massif, winter, fell, ridge, geology, hill station, plateau, glacial landform, shield volcano, freezing, wilderness, extinct volcano, cinder cone, summit, lava dome, evening, forest, lake district, volcanic landform, sunrise, cumulus, street light, loch, reservoir, ice cap, dusk, reflection

ตารางที่ 24 เมทาดาทาจากการทดลองในหมวดหมู่ภาพวิวทิวทัศน์ รหัสภาพ VL07

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ VL07	7	boardwalks, grass, mountain, path, ramp, walk, walkway
อเมซอน เรคคอคนิชัน	18	boardwalk, building, bridge, person, human, path, outdoors, nature, trail, scenery, grass, plant, walkway, landscape, plateau, mountain range, mountain
คลาริฟาย	21	landscape, mountain, travel, nature, sky, grass, no person, hike, guidance, lake, snow, outdoors, trail, summer, valley, water, hill, adventure, road, fog
อิมเมกกา	10	ascent, highland, slope, mountain, landscape, mountains, sky, travel, grass, scenery
กูเกิลคลาวน์วิชันเอพีไอ	46	cloud, sky, mountain, plant, natural landscape, slope, outdoor recreation, terrain, people in nature, grassland, road, grass, landscape, recreation, mountain range, leisure, travel, valley, hiking, hill, adventure, plateau, fell, trail, lake district, walking, ridge, massif, pasture, path, forest, reservoir, glacial landform, tundra, summit, mountain pass, walkway, vacation

ตารางที่ 25 เมทาดาทาจากการทดลองในหมวดหมู่ภาพวิวทิวทัศน์ รหัสภาพ VL09

เครื่องมือ	จำนวนคำ	เมทาดาทา
 รหัสภาพ VL09	6	air sport, blue, glider, gliding, parachute, sky
อเมซอน เรคคอคนิชัน	6	adventure, leisure activities, gliding, parachute, person, human
คลาริฟาย	20	mountain, snow, adventure, landscape, parachute, high, sky, paraglider, travel, air, flying, recreation, winter, altitude, climb, sport, glider, freedom, valley, nature
อิมเมกกา	10	parachute, rescue equipment, equipment, mountain, snow, sky, mountains, landscape, peak, extreme
กูเกิลคลาวน์วิชันเอพีไอ	37	cloud, sky, mountain, parachute, paragliding, snow, slope, parachuting, cumulus, wind sports, geological phenomenon, landscape, wind, air sports, sports, recreation, travel, air travel, mountain range, leisure, horizon, hill, extreme sport, hill station, fell, meteorological phenomenon, massif, adventure, fun, winter, summit, grassland, valley, glacial landform

หมวดหมู่ภาพผัก ผลไม้

ตารางที่ 26 เมทาดาทาจากการทดลองในหมวดหมู่ภาพผัก ผลไม้ รหัสภาพ VF04

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ VF04	6	cooking, food, produce, radish, red, vegetables
อเมซอน เรคคอคนิชัน	5	radish, plant, vegetable, food, produce
คลาริฟาย	20	food, health, no person, pasture, nutrition, radish, leaf, healthy, farming, vegetable, grow, nature, garden, juicy, diet, fruit, vitamin, bunch, delicious, market
อิมเมกกา	10	radish, root vegetable, radish plant, fruit, vegetable, ripe, food, healthy, sweet, herb
กูเกิลคลาวน์วิชันเอพีไอ	44	food, plant, fruit, leaf, ingredient, natural foods, red, staple food, superfood, whole food, produce, flowering plant, rose family, alpine strawberry, accessory fruit, local food, pear, berry, petal, groundcover, frutti di bosco, superfruit, rose order, apple, vegan nutrition, event, tree, carmine, vegetable, mcintosh, still life photography, fruit tree, pear, radish, sweetness, annual plant, photography, arctostaphylos, subshrub

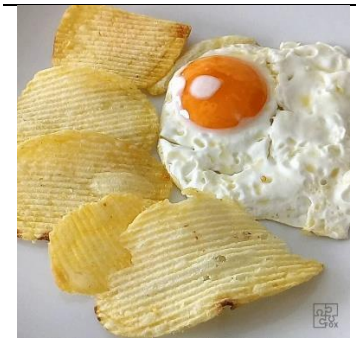
ตารางที่ 27 เมทาดาทาจากการทดลองในหมวดหมู่ภาพผัก ผลไม้ รหัสภาพ VF05

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ VF05	6	apples, food, fruit, grapes, oranges, plums
อเมซอน เรคคอคนิชัน	9	plant, fruit, food, produce, grapes, plum, pineapple, orange, citrus fruit
คลาริฟาย	21	fruit, market, juicy, food, grow, abundance, confection, many, health, plum, group, nutrition, no person, stall, pasture, kind, apple, pineapple, healthy, pear
อิมเมกกา	11	fruit, pumpkin, edible fruit, produce, food, orange, vegetable, vitamin, citrus, apple
กูเกิลคลาวน์วิชันเอพีไอ	55	food, valencia orange, clementine, rangpur, fruit, ingredient, plant, tangerine, natural foods, bitter orange, seedless fruit, orange, tangelo, grapefruit, mandarin orange, citrus, calamondin, orange, whole food, cuisine, food group, market, greengrocer, public space, superfood, produce, citric acid, local food, ananas, marketplace, human settlement, sweetness, accessory fruit, vegan nutrition, basket, city, grocery store, bazaar, still life photography, lime, delicacy, still life, retail, landscape, superfruit, pomelo, recipe, vegetarian food, pattern

ตารางที่ 28 เมทาดาทาจากการทดลองในหมวดหมู่ภาพผัก ผลไม้ รหัสภาพ VF06

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ VF06	7	chard, farmer, lettuce, local, market, produce, vegetables,
อเมซอน เรคคอคนิชัน	8	person, human, plant, spinach, vegetable, food, shop, produce
คลาริฟาย	18	lettuce, food, ingredients, market, kind, nutrition, nature, leaf, grow, cabbage, shopping, vitamin, pasture, vegetable, cooking, basket, dinner, delicious
อิมเมกกา	9	lettuce, greens, spinach, fresh, herb, food, salad, produce, leaf,
กูเกิลคลาวน์วิชันเอพีไอ	37	ingredient, leaf vegetable, whole food, public space, handwriting, vegetable, market, staple food, local food, produce, basket, human settlement, superfood, greengrocer, marketplace, cuisine, vegan nutrition, cruciferous vegetables, dish, retail, city, food group, storage basket, herb, flowering plant, trade, comfort food, wild cabbage, cabbage, food storage containers, iceberg lettuce

ตารางที่ 29 เมทาดาทาจากการทดลองในหมวดหมู่ภาพอาหาร รหัสภาพ F07

เครื่องมือ	จำนวนคำ	เมทาดาทา
 รหัสภาพ F07	6	breakfast, chip, egg, food, meal, snack
อเมซอน เรคคอคนิชัน	6	egg, food, meal, dish, sweets, confectionery
คลาริฟาย	20	no person, food, crisp, breakfast, nutrition, delicious, cooking, gold, butter, health, grow, egg, meal, corn, ingredients, lunch, dawn, traditional, homemade, unhealthy
อิมเมกกา	10	cheese, food, meal, plate, snack, breakfast, delicious, slice, dessert, fresh
กูเกิลคลาวน์วิชันเอพีไอ	40	food, tableware, fried egg, egg yolk, egg white, ingredient, cuisine, egg, recipe, dish, baked goods, produce, staple food, wood, junk food, dishware, comfort food, finger food, blue cheese dressing, fast food, dairy, fried food, frying, poached egg, white bread, breakfast, side dish, gluten, plate, platter, meal, mayonnaise, condiment, cheese spread, meat, vegan nutrition, banana, spoon

ตารางที่ 30 เมทาดาทาจากการทดลองในหมวดหมู่ภาพอาหาร รหัสภาพ F09

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ F09	6	cheese, dish, food, italian, pizza, slice
อเมซอน เรคคอคนิชัน	2	pizza, food
คลาริฟาย	20	pizza, cheese, meal, lunch, food, dinner, tomato, slice, hot, vegetable, mushroom, meat, mozzarella, sauce, crust, restaurant, tasty, refreshment, dough, delicious
อิมเมกกา	10	pizza, dish, food, nutriment, cheese, tomato, dinner, meal, lunch, sauce
กูเกิลคลาวน์วิชันเอพีไอ	43	food, pizza, ingredient, recipe, fast food, cuisine, california-style pizza, pizza cheese, baked goods, dish, produce, staple food, leaf vegetable, junk food, finger food, vegetable, comfort food, processed cheese, meat, flatbread, cheese, take-out food, garnish, cooking, italian food, georgian cuisine, romano cheese, breakfast, side dish, american food, dairy, baking, meal, french food, culinary art, frittata, vegan nutrition, brunch, indian cuisine, mexican food, vegetarian food, tarte flambée, parmigiano-reggiano

ตารางที่ 31 เมทาดาทาจากการทดลองในหมวดหมู่ภาพอาหาร รหัสภาพ F15

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ F15	6	biscuit, cake, decorating, icing, train, sweet
อเมซอน เรคคอคนิชัน	12	icing, cream, dessert, cake, food, creme, sweets, confectionery, birthday cake, cookie, biscuit, gingerbread
คลาริฟาย	20	no person, cake, chocolate, sugar, candy, baking, delicious, sweet, cream, sprinkles, birthday, party, cookie, indulgence, food, celebration, traditional, unhealthy, wood, homemade
อิมเมกกา	10	pillow, cushion, padding, decoration, celebration, birthday, box, cake, gift, dessert
กูเกิลคลาวน์วิชันเอพีไอ	51	food, cake decorating, cake, cake decorating supply, ingredient, recipe, birthday party, baked goods, pink, cuisine, gingerbread house, sugar paste, sugar cake, cream, icing, birthday cake, birthday, buttercream, wood, sweetness, fondant, event, dessert, baking, kuchen, marshmallow creme, dish, finger food, house, tableware, confectionery, whipped cream, party, christmas decoration, table, snack, pastry, gingerbread, plate, interior design, royal icing, torte, pasteles, frozen dessert, dairy, culinary art

หมวดหมู่ภาพยานพาหนะ

ตารางที่ 32 เมทาดาทาจากการทดลองในหมวดหมู่ภาพยานพาหนะ รหัสภาพ V03

เครื่องมือ	จำนวนคำ	เมทาดาทา
 รหัสภาพ V03	6	car, colorful, minibus, tree, thailand, van
อเมซอน เรคคอคนิชัน	13	bus, vehicle, transportation, plant, van, minibus, truck, wheel, machine, flower, blossom, tarmac, asphalt
คลาริฟาย	21	car, vehicle, bus, traffic, transportation system, road, street, truck, city, drive, driver, wheel, police, landscape, accident, automotive, emergency, travel, people, machine
อิมเมกกา	11	minibus, bus, car, public transport, transportation, road, vehicle, transport, truck, conveyance
กูเกิลคลาวน์วิชันเอพีไอ	29	tire, wheel, land vehicle, car, vehicle, motor vehicle, window, plant, mode of transport, yellow, van, light commercial vehicle, automotive wheel system, tree, commercial vehicle, compact van, automotive exterior, street, vehicle registration plate, public transport, transport, magenta, city car, minibus, microvan, automotive parking light, minivan

ตารางที่ 33 เมทาดาทาจากการทดลองในหมวดหมู่ภาพยานพาหนะ รหัสภาพ V07

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ V07	7	austria, biker, boots, freedom, harley, motorcycle, touring
อเมซอน เรคคอคนิชัน	14	motorcycle, transportation, vehicle, person, human, machine, clothing, apparel, helmet, motor, crash helmet, hardhat, engine, spoke
คลาริฟาย	22	bike, street, road, wheel, vehicle, motorbike, biker, city, people, transportation system, seated, hurry, man, cyclist, rider, drive, driver, urban, roll along, woman
อิมเมกกา	11	helmet, motorcycle, crash helmet, bike, speed, motor, vehicle, sidecar, motorbike, conveyance
กูเกิลคลาวน์วิชันเอพีไอ	38	tire, jeans, wheel, fuel tank, vehicle, motor vehicle, automotive lighting, automotive tire, leg, motorcycle, sunglasses, automotive design, helmet, mode of transport, motorcycle accessories, automotive wheel system, rim, fender, motorcycling, headlamp, motorcycle helmet, automotive exterior, road, vehicle brake, spoke, cruiser, auto part, classic, electric blue, personal protective equipment, street, eyewear, brake, chopper, transport

ตารางที่ 34 เมทาดาทาจากการทดลองในหมวดหมู่ภาพยานพาหนะ รหัสภาพ V10

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ V10	6	bus, bus stop, london, people, road, route177
อเมซอน เรคคอคนิชัน	14	bus, vehicle, transportation, person, human, tour bus, tarmac, asphalt, double decker bus, bus stop, road, pedestrian
คลาริฟาย	20	street, city, bus, tramway, road, travel, urban, traffic, tram, people, public, tourist, locomotive, architecture, vehicle, transportation system, railway, shopping, sightseeing, business
อิมเมกกา	11	tramway, conveyance, vehicle, streetcar, transportation, transport, car, road, wheeled vehicle, street
กูเกิลคลาวน์วิชันเอพีไอ	32	bus, sky, vehicle, building, double-decker bus, motor vehicle, window, tour bus service, waste container, vehicle registration plate, tire, travel, wheel, city, public transport, asphalt, road, commercial vehicle, electricity, pedestrian, mixed-use, street, traffic, lane, public utility, downtown, cloud, automotive exterior, advertising, passenger, compact car, street light, transport, tourism

ตารางที่ 35 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสถานที่ท่องเที่ยว รหัสภาพ L01

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ L01	9	london eye, sunny, ferris, Thames, England, tourists, attraction
อเมซอน เรคคอคนิชัน	13	ferris wheel, amusement park, boat, transportation, vehicle, waterfront, water, city, building, town, urban
คลาริฟาย	21	water, river, bridge, travel, sunset, architecture, city, sky, reflection, cityscape, urban, no person, dawn, building, boat, ferris wheel, landmark, tourism, tower, dusk
อิมเมกกา	10	park, city, bridge, river, tract, waterfront, architecture, water, sky, boat
กูเกิลคลาวน์วิชันเอพีไอ	44	cloud, water, sky, building, wheel, ferris wheel, lighting, dusk, lake, plant, tower block, city, cityscape, morning, tree, boat, watercraft, landmark, leisure, metropolitan area, recreation, metropolis, fun, amusement ride, circle, horizon, reflection, evening, event, night, automotive wheel system, nonbuilding structure, tourism, tower, town, amusement park, arch, skyline, landscape, cumulus, downtown, tourist attraction, channel, roller coaster, water transportation

ตารางที่ 36 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสถานที่ท่องเที่ยว รหัสภาพ L06

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ L06	7	architecture, leaning, leaning tower of pisa, pisa, sky, tower, outdoor
อเมซอน เรคคอคนิชัน	9	building, architecture, tower, monument, urban, brick, city, town, metropolis
คลาริฟาย	20	architecture, pisa, travel, sky, no person, old, tower, ancient, renaissance, outdoors, marble, building, city, landmark, art, plaza, famous, monument, gothic, traditional
อิมเมกกา	10	architecture, building, tower, dome, church, landmark, cathedral, history, city, facade
กูเกิลคลาวน์วิชันเอพีไอ	27	cloud, sky, building, tower, facade, monument, city, column, medieval architecture, symmetry, cylinder, ancient history, history, ancient roman architecture, historic site, classical architecture, engineering, tower block, national historic landmark, art, tree, composite material, unesco world heritage site, tourist attraction, observation tower, baptistery

ตารางที่ 37 เมทาดาทาจากการทดลองในหมวดหมู่ภาพสถานที่ท่องเที่ยว รหัสภาพ L15

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ L15	7	ancient, egypt, giza, old kingdom, pyramid, 4th dynasty, stone
อเมซอน เรคคอคนิชัน	5	architecture, building, person, human, pyramid
คลาริฟาย	20	pyramid, no person, pharaoh, desert, sphinx, grave, archaeology, ancient, travel, sand, architecture, camel, stone, mausoleum, people, outdoors, sky, religion, temple, step
อิมเมกกา	10	grave, tourism, desert, travel, sky, stone, ancient, architecture, history, pyramid
กูเกิลคลาวน์วิชันเอพีไอ	23	sky, pyramid, wonders of the world, landscape, travel, monument, archaeological site, tints and shades, triangle, ancient history, history, sand, aeolian landform, art, historic site, working animal, symmetry, tourist attraction, tower, unesco world heritage site, paper product, national monument, rock

หมวดหมู่ภาพศิลปวัฒนธรรม

ตารางที่ 38 เมทาดาทาจากการทดลองในหมวดหมู่ภาพศิลปวัฒนธรรม รหัสภาพ AC05

เครื่องมือ	จำนวนคำ	เมทาดาทา
รหัสภาพ AC05	6	thailand, bangkok, grand palace, giant, temple, palace
อเมซอน เรคคอคนิชัน	16	architecture, building, temple, person, human, shrine, worship, roof, spire, tower, steeple, pillar, column, monastery, housing, pagoda
คลาริฟาย	20	temple, buddha, culture, religion, wat, travel, art, sculpture, castle, traditional, statue, gold, monastery, pagoda, architecture, shrine, worship, decoration, grand, roof
อิมเมกกา	10	temple, religion, architecture, culture, gold, palace, religious, east, traditional, travel
กูเกิลคลาวน์วิชันเอพีไอ	24	building, temple, statue, sky, sculpture, chinese architecture, temple, facade, art, leisure, japanese architecture, place of worship, event, historic site, holy places, hindu temple, tourism, tradition, monastery, roof, tourist attraction, shrine, religious institute, visual arts

ตารางที่ 39 เมทาดาทาจากการทดลองในหมวดหมู่ภาพศิลปวัฒนธรรม รหัสภาพ AC10

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ AC10	5	buddha, monk, temple, travel, thailand
อเมซอน เรคคอคนิชัน	11	person, human, architecture, building, monk, temple, shrine, worship, people, monastery, housing
คลาริฟาย	20	temple, buddha, religion, travel, monk, people, worship, shrine, monastery, architecture, religious, tourism, culture, spirituality, gold, sculpture, statue, prayer, city, pagoda
อิมเมกกา	10	monk, temple, religion, architecture, travel, religious, building, statue, culture, old
กูเกิลคลาวน์วิชันเอพีไอ	27	world, temple, building, leisure, temple, travel, landmark, sangharaja, city, art, pilgrimage, event, holy places, place of worship, religious institute, tourism, monk, chinese architecture, worship, reflection, ritual, hindu temple, historic site, shrine, tourist attraction, pedestrian, monastery, night

ตารางที่ 40 เมทาดาทาจากการทดลองในหมวดหมู่ภาพศิลปวัฒนธรรม รหัสภาพ AC30

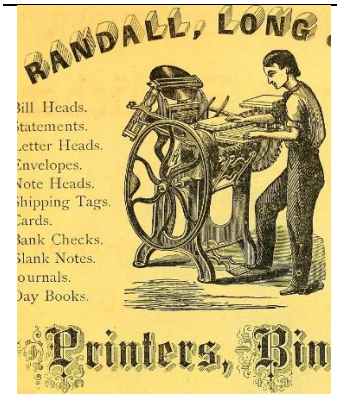
เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ AC30	7	france, louis xiv, palace, paris, sculpture, versailles, water garden
อเมซอน เรคคอคนิชัน	7	statue, sculpture, art, person, human, monument, water
คลาริฟาย	20	sculpture, statue, art, travel, fountain, monument, sky, water, bronze, architecture, city, culture, old, tourism, symbol, ancient, religion, park, marble, people
อิมเมกกา	10	statue, fountain, structure, sculpture, architecture, monument, history, bronze, culture, landmark
กูเกิลคลาวน์วิชันเอพีไอ	34	water, sky, cloud, plant, eye, sculpture, statue, world, tree, art, fountain, leisure, monument, classical sculpture, water feature, metal, city, mythology, nonbuilding structure, garden, bronze sculpture, memorial, artwork, town square, national historic landmark, tourism, bronze, vacation, visual arts, sitting, stone carving, grass

หมวดหมู่ภาพปกหนังสือ/โปสเตอร์เก่า

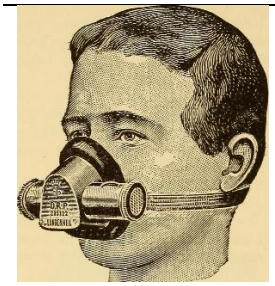
ตารางที่ 41 เมทาดาทาจากการทดลองในหมวดหมู่ภาพปกหนังสือ/โปสเตอร์เก่า รหัสภาพ BP02

เครื่องมือ	**จำนวนคำ**	**เมทาดาทา**
 รหัสภาพ BP02	5	bookid:bostoncookingsch19hill_4, bookyear:1896, bookdecade:1890, bookcentury:1800, bookauthor:Hill__Janet_McKenzi
อเมซอน เรคคอคนิชัน	14	furniture, person, human, chair, text, advertisement, poster, clothing, apparel, sitting, shoe, footwear, page, art
คลาริฟาย	21	people, one, retro, child, sit, adult, education, text, nostalgia, administration, two, antique, wear, vintage, furniture, monochrome, man, newspaper, portrait, book series
อิมเมกกา	10	newspaper, product, creation, daily, man, old, male, vintage, people, person
กูเกิลคลาวน์วิชันเอพีไอ	23	poster, font, baby, vintage clothing, illustration, toddler, chair, child, photo caption, book cover, happy, advertising, sitting, vintage advertisement, pattern, retro style, baby & toddler clothing, publication, art

ตารางที่ 42 เมทาดาทาจากการทดลองในหมวดหมู่ภาพปกหนังสือ/โปสเตอร์เก่า รหัสภาพ BP03

เครื่องมือ	จำนวนคำ	เมทาดาทา
 รหัสภาพ BP03	7	bookpublisher:_New_York__American_Telephone_and_Telegraph_Co___etc__, bookcontributor:Prelinger_Library, booksponsor:Internet_Archive, bookleafnumber:399, bookcollection:prelinger_library, bookcollection:additional_collections, bookcollection:americana
อเมซอน เรคคอคนิชัน	19	person, woman, adult, female, advertisement, wheel, machine, poster, face, head, text, book, publication, sewing, chair, furniture, painting, art, page,
คลาริฟาย	23	illustration, art, vintage, people, retro, print, chalk out, old, paper, antique, woodcut, man, ancient, bill, sketchy, book bindings, lithograph, painting, engraving, sepia pigment
อิมเมกกา	10	treasury, depository, vintage, facility, retro, paper, old, drawing, design, art
กูเกิลคลาวน์วิชันเอพีไอ	15	poster, font, art, illustration, drawing, line art, motor vehicle, vintage advertisement, history, visual arts, human anatomy, paper product, circle, printmaking, machine

ตารางที่ 43 เมทาดาทาจากการทดลองในหมวดหมู่ภาพปกหนังสือ/โปสเตอร์เก่า รหัสภาพ BP07

เครื่องมือ	จำนวนคำ	เมทาดาทา
 รหัสภาพ BP07	7	booksubject:Sanitation, bookpublisher:New_York__Pub__for_Sears__Roebuck_and_co__by_D__Appleton_and_company, bookcontributor:The_Library_of_Congress, booksponsor:The_Library_of_Congress, bookleafnumber:961, bookcollection:library_of_congress, bookcollection:americana
อเมซอน เรคคอคนิชัน	10	person, man, adult, male, art, face, head, drawing, photography, painting
คลาริฟาย	20	illustration, man, people, portrait, old, vintage, art, head, one, print, soldier, retro, war, military, face, antique, wear, warrior, money, ancient
อิมเมกกา	13	chain mail, covering, mask, body armor, armor, protective covering, disguise, attire, face, clothing
กูเกิลคลาวน์วิชันเอพีไอ	27	nose, jaw, gesture, art, hearing, illustration, personal protective equipment, wrist, beard, eyewear, drawing, facial hair, bottle, audio equipment, font, line art, drinkware, drinking, painting, circle, self-portrait, gas mask, sketch, visual arts, sports gear, history, artwork

ผลการคำนวณค่าความถูกต้องเมทาดาทา และประสิทธิภาพเครื่องมือ

การวิเคราะห์ประสิทธิภาพของเครื่องมือการรู้จำภาพถ่ายนี้ใช้วิธีการทั้งเชิงปริมาณและคุณภาพ ในเชิงปริมาณจะตรวจสอบจำนวนภาพถ่ายที่สามารถแยกเมทาดาทาได้ค่าความถูกต้องมากกว่า 80% ในเชิงคุณภาพใช้วิธีการวิเคราะห์ข้อมูลสองแบบ คือ วิธีการทางสถิติเพื่อหาค่าความคล้ายคลึงของเอกสาร (Similarity) ได้แก่ การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure) การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap SimilarityMeasure) การวัดค่าความคล้ายโคไซน์ (Cosine Similarity) และวิธีวัดประสิทธิภาพเครื่องมือได้แก่ การหาค่าความไว (Recall) ค่าความแม่นยำ (Precision) และ ค่าความถูกต้องเอฟวัน (F1-score) โดยแยกตามหมวดหมู่ภาพ

ตัวอย่างผลการวิเคราะห์เมทาดาทาในภาพหมวดหมู่คน

ตารางที่ 44 เมทาดาทาเรียงตามลำดับตัวอักษรจากตัวอย่างภาพคน

เครื่องมือ	จำนวนคำ	เมทาดาทาเรียงตามตัวอักษร
 รหัสภาพ H01	7	action, alpine, Korea, skiing, snow, sport, winter
อเมซอน เรคคอคนิชัน	20	aerial view, architecture, azure sky, building, city, downtown, high rise, landscape, metropolis, nature, neighborhood, outdoors, scenery, sky, spire, steeple, tower, town, urban
คลาริฟาย	20	action, action energy, alpine, cold, competition, downhill, fast, fun, gloves, goggles, helmet, hurry, ice, man, race, resort, skier, snow, sport, winter
อิมเมกกา	10	cold, fun, mountain, outdoor, season, ski, skier, snow, sport, winter
กูเกิลคลาวน์วิชันเอพีไอ	48	competition, competition event, cross-country skiing, exercise, eyewear, footwear, freezing, fun, glacial landform, glove, goggles, hat, headgear, helmet, ice cap, individual sports,

เครื่องมือ	จำนวนคำ	เมทาดาทาเรียงตามตัวอักษร
		jumping, leisure, Nordic combined, Nordic skiing, outdoor recreation, personal protective equipment, piste, playing in the snow, recreation, shoe, ski, ski binding, ski boot, ski cross, ski equipment, ski helmet, ski pole, ski touring, skier, skiing, slalom skiing, slope, sneakers, snow, sports, sports equipment, sports gear, sports uniform, sportswear, trousers, white, winter, winter sport

ตารางที่ 45 เมทาดาทาทับซ้อนของตัวอย่างภาพคน

เครื่องมือ	จำนวนคำ	เมทาดาทาทับซ้อน
อเมซอน เรคคอคนิชัน	0	-
คลาริฟาย	6	action, alpine, skier, snow, sport, winter
อิมเมกกา	5	ski, skier, snow, sport, winter
กูเกิลคลาวน์วิชันเอพีไอ	4	ski, snow, sport, winter, winter sport

ตารางที่ 46 เมทาดาทาที่มีความหมายเกี่ยวข้องของตัวอย่างภาพคน

เครื่องมือ	จำนวนคำ	เมทาดาทาที่มีความหมายเกี่ยวข้อง
อเมซอน เรคคอคนิชัน	3	outdoor, scenery, steeple
คลาริฟาย	5	cold, gloves, goggles, helmet, race,
อิมเมกกา	2	cold, outdoor
กูเกิลคลาวน์วิชันเอพีไอ	10	competition, exercise, eyewear, glove, googles, helmet, individual sports, ski equipment, sportswear, slope

จากนั้นนำจำนวนคำที่วิเคราะห์ได้ทั้งเมทาดาทาที่ทับซ้อน และเมทาดาทาที่มีความหมายเกี่ยวข้องไปคำนวณค่าทางสถิติต่าง ๆ โดยแยกผลการทดลองตามหมวดหมู่ภาพ

หมวดภาพคน

ตารางที่ 47 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพคน

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	70.15	60.00	50.78
คลาริฟาย	**77.25**	80.00	**75.56**
อิมเมกกา	70.11	90.00	45.00
กูเกิลคลาวน์วิชันเอพีไอ	60.02	**120.00**	35.00

ตารางที่ 48 ค่าประสิทธิภาพของเมทาดาทาภาพในหมวดภาพคน

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	68.00	70.00	68.99
คลาริฟาย	85.56	77.00	**81.05**
อิมเมกกา	57.00	60.00	58.46
กูเกิลคลาวน์วิชันเอพีไอ	55.00	59.00	56.93

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่ โปรแกรมคลาริฟาย อเมซอน เรคคอคนิชัน อิมเมกกา และ กูเกิลคลาวน์วิชันเอพีไอ ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมคลาริฟายมีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 70.25% 75.56% และ 81.05% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 120%

หมวดภาพสัตว์

ตารางที่ 49 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพสัตว์

เครื่องมือ	การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)	การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)	การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)
อเมซอน เรคคอคนิชัน	71.44	60.00	54.26
คลาริฟาย	**72.32**	80.00	**74.56**
อิมเมกกา	69.55	90.00	48.57
กูเกิลคลาวน์วิชันเอพีไอ	62.11	**100.00**	37.00

ตารางที่ 50 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพสัตว์

เครื่องมือ	ค่าความไว (Recall)	ค่าความแม่นยำ (Precision)	ค่าความถูกต้องเอฟวัน (F1-score)
อเมซอน เรคคอคนิชัน	65.00	73.00	68.77
คลาริฟาย	88.56	80.00	**84.06**
อิมเมกกา	55.00	63.00	58.73
กูเกิลคลาวน์วิชันเอพีไอ	50.00	61.00	54.95

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่ โปรแกรมคลาริฟาย อเมซอน เรคคอคนิชัน อิมเมกกา และ กูเกิลคลาวน์วิชันเอพีไอ ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมคลาริฟายมีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 72.32% 74.56% และ 84.06% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 100%

หมวดภาพต้นไม้ ดอกไม้

ตารางที่ 51 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพต้นไม้ ดอกไม้

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	50.11	80.00	48.56
คลาริฟาย	64.78	100.00	56.54
อิมเมกกา	**75.77**	110.00	**75.50**
กูเกิลคลาวน์วิชันเอพีไอ	66.35	**120.00**	57.78

ตารางที่ 52 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพต้นไม้ ดอกไม้

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	67.00	50.00	57.26
คลาริฟาย	56.00	75.00	64.12
อิมเมกกา	77.56	89.12	**82.94**
กูเกิลคลาวน์วิชันเอพีไอ	60.00	71.00	65.04

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรมอิมเมกกา กูเกิลคลาวน์วิชันเอพีไอ คลาริฟาย และ อเมซอน เรคคอคนิชัน ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมอิมเมกกามีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 75.77% 75.50% และ 82.94% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 120%

หมวดภาพวิวทิวทัศน์

ตารางที่ 53 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพวิวทิวทัศน์

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	72.15	70.00	57.67
คลาริฟาย	71.11	100.00	56.75
อิมเมกกา	61.24	110.00	45.24
กูเกิลคลาวน์วิชันเอพีไอ	**74.48**	**120.00**	**65.66**

ตารางที่ 54 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพวิวทิวทัศน์

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	66.00	78.00	71.50
คลาริฟาย	65.00	76.00	70.07
อิมเมกกา	64.00	55.00	59.16
กูเกิลคลาวน์วิชันเอพีไอ	78.00	87.57	**82.51**

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อเมซอน เรคคอคนิชัน คลาริฟาย และอิมเมกกา ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมกูเกิลคลาวน์วิชันเอพีไอมีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 74.48% 65.66% และ 82.51% ตามลำดับ และมีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 120%

หมวดภาพผัก ผลไม้

ตารางที่ 55 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพผัก ผลไม้

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	**78.23**	70.00	**84.87**
คลาริฟาย	60.11	80.00	56.54
อิมเมกกา	77.31	85.00	75.10
กูเกิลคลาวน์วิชันเอพีไอ	50.14	**90.00**	48.56

ตารางที่ 56 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพผัก ผลไม้

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	89.24	80.00	**84.37**
คลาริฟาย	57.00	56.00	56.50
อิมเมกกา	74.00	81.00	77.34
กูเกิลคลาวน์วิชันเอพีไอ	40.00	74.00	51.93

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรมอเมซอน เรคคอคนิชัน อิมเมกกา คลาริฟาย และกูเกิลคลาวน์วิชันเอพีไอ ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมอเมซอน เรคคอคนิชันมีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 78.23% 84.87% และ 84.37% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 90%

หมวดภาพอาหาร

ตารางที่ 57 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพอาหาร

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	49.56	60.00	45.58
คลาริฟาย	**73.25**	80.00	**85.47**
อิมเมกกา	55.21	100.00	51.28
กูเกิลคลาวน์วิชันเอพีไอ	61.02	**120.00**	65.10

ตารางที่ 58 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพอาหาร

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	65.00	72.00	68.32
คลาริฟาย	87.59	80.00	**83.62**
อิมเมกกา	65.00	78.00	70.91
กูเกิลคลาวน์วิชันเอพีไอ	74.00	82.00	77.79

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรมคลาริฟาย กูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา และอเมซอน เรคคอคนิชัน ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมคลาริฟายมีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 73.25% 85.47% และ 83.62% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 120%

หมวดภาพยานพาหนะ

ตารางที่ 59 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพยานพาหนะ

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	**77.79**	65.00	**76.87**
คลาริฟาย	55.50	100.00	35.40
อิมเมกกา	64.02	110.00	57.88
กูเกิลคลาวน์วิชันเอพีไอ	36.55	**120.00**	31.29

ตารางที่ 60 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพยานพาหนะ

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	84.52	83.00	**83.75**
คลาริฟาย	65.00	76.00	70.07
อิมเมกกา	75.00	83.00	78.80
กูเกิลคลาวน์วิชันเอพีไอ	65.00	71.00	67.87

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรม อเมซอน เรคคอคนิชัน อิมเมกกา คลาริฟาย และกูเกิลคลาวน์วิชันเอพีไอ ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมอเมซอน เรคคอคนิชันมีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 77.79% 76.87% และ 83.75% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 120%

หมวดภาพสถานที่ท่องเที่ยว

ตารางที่ 61 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพสถานที่ท่องเที่ยว

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	71.24	88.00	55.01
คลาริฟาย	75.35	90.00	59.56
อิมเมกกา	70.45	95.00	51.22
กูเกิลคลาวน์วิชันเอพีไอ	**77.10**	**100.00**	**75.88**

ตารางที่ 62 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพสถานที่ท่องเที่ยว

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	63.00	55.00	58.73
คลาริฟาย	86.00	76.00	80.69
อิมเมกกา	58.00	56.00	56.98
กูเกิลคลาวน์วิชันเอพีไอ	84.56	81.24	**82.87**

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ คลาริฟาย อเมซอน เรคคอคนิชัน และอิมเมกกา ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 77.10% 75.88% และ 82.87% ตามลำดับ และมีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 100%

หมวดภาพศิลปวัฒนธรรม

ตารางที่ 63 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพศิลปวัฒนธรรม

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	61.47	78.00	64.23
คลาริฟาย	68.89	85.00	74.15
อิมเมกกา	**78.45**	90.00	**81.51**
กูเกิลคลาวน์วิชันเอพีไอ	56.64	**120.00**	47.21

ตารางที่ 64 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพศิลปวัฒนธรรม

เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	76.00	65.00	70.07
คลาริฟาย	75.00	86.00	80.12
อิมเมกกา	84.00	85.00	**84.50**
กูเกิลคลาวน์วิชันเอพีไอ	54.00	65.00	58.99

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรมอิมเมกกา คลาริฟาย อเมซอน เรคคอคนิชัน และกูเกิลคลาวน์วิชันเอพีไอ ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมอิมเมกกามีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 78.45% 81.51% และ 84.50% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 120%

หมวดภาพปกหนังสือ/โปสเตอร์เก่า

ตารางที่ 65 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาในหมวดภาพปกหนังสือ/โปสเตอร์เก่า

เครื่องมือ	การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)	การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)	การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)
อเมซอน เรคคอคนิชัน	69.12	80.00	49.56
คลาริฟาย	62.58	85.00	40.23
อิมเมกกา	**78.23**	90.00	**74.00**
กูเกิลคลาวน์วิชันเอพีไอ	71.07	**150.00**	62.71

ตารางที่ 66 ค่าประสิทธิภาพเฉลี่ยของเมทาดาทาในหมวดภาพปกหนังสือ/โปสเตอร์เก่า

เครื่องมือ	ค่าความไว (Recall)	ค่าความแม่นยำ (Precision)	ค่าความถูกต้องเอฟวัน (F1-score)
อเมซอน เรคคอคนิชัน	68.00	74.00	70.87
คลาริฟาย	58.00	66.00	61.74
อิมเมกกา	83.69	83.69	**89.00**
กูเกิลคลาวน์วิชันเอพีไอ	80.00	84.00	81.95

ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน เรียงตามลำดับจากมากไปน้อย ได้แก่โปรแกรมอิมเมกกา กูเกิลคลาวน์วิชันเอพีไอ อเมซอน เรคคอคนิชัน และคลาริฟาย ในขณะที่ค่าคำที่ไม่ทับซ้อนเรียงจากมากไปน้อยได้แก่ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ อิมเมกกา คลาริฟาย และอเมซอน เรคคอคนิชัน ตามลำดับ

โปรแกรมอิมเมกกามีค่าคำทับซ้อน ค่าความคล้ายโคไซน์ ค่าความถูกต้องเอฟวันสูงที่สุดเท่ากับ 78.23% 74.00% และ 89.00% ตามลำดับ ในขณะที่โปรแกรมกูเกลิคลาวน์วิชันเอพีไอ มีค่าคำที่ไม่ทับซ้อนแต่เป็นคำที่ให้ความหมายที่เกี่ยวข้องเหมาะสม มีโอกาสในการค้นหาเพิ่มขึ้นสูงสุดที่ 150%

ตารางที่ 67 ค่าความคล้ายคลึงเฉลี่ยของเมทาดาทาจากการทดลองทั้งหมด 1,000 ภาพ

เครื่องมือ	**การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure)**	**การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure)**	**การวัดค่าความคล้ายโคไซน์ (Cosine Similarity)**
อเมซอน เรคคอคนิชัน	57.84	73.60	58.64
คลาริฟาย	67.41	87.50	59.48
อิมเมกกา	70.03	87.50	59.48
กูเกิลคลาวน์วิชันเอพีไอ	58.55	116.00	47.62

ตารางที่ 68 ค่าประสิทธิภาพเฉลี่ยจากการทดลองภาพทั้งหมด 1,000 ภาพ

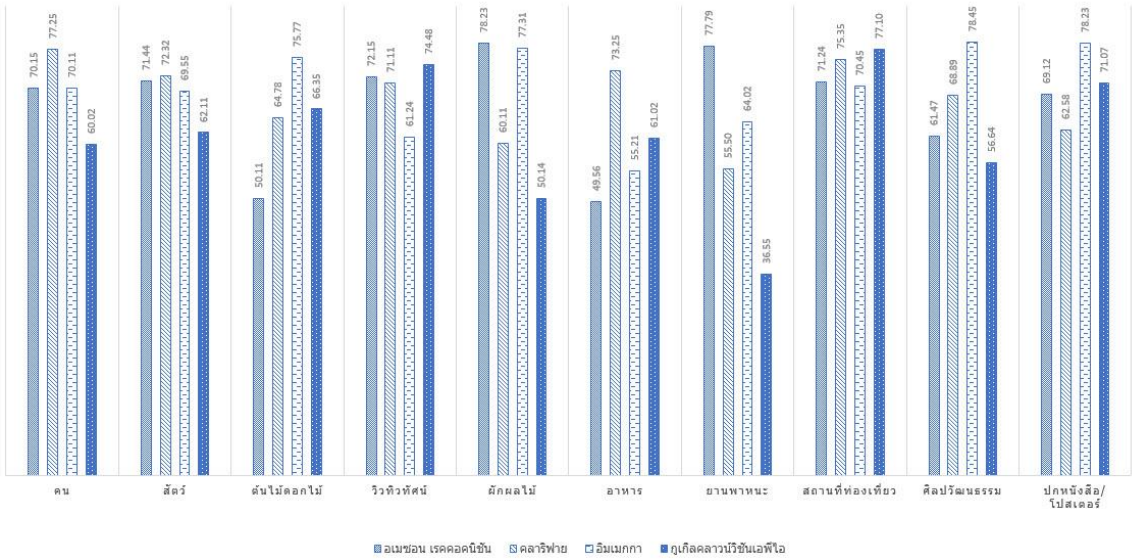
เครื่องมือ	**ค่าความไว (Recall)**	**ค่าความแม่นยำ (Precision)**	**ค่าความถูกต้องเอฟวัน (F1-score)**
อเมซอน เรคคอคนิชัน	63.34	74.70	64.21
คลาริฟาย	68.77	79.90	65.06
อิมเมกกา	67.39	80.31	64.25
กูเกิลคลาวน์วิชันเอพีไอ	56.54	96.26	56.36

พิจารณาผลรวมเฉลี่ยของทุกภาพแล้วพบว่าค่าความคล้ายคลึง และความถูกต้องเอฟวันของเมทาดาทาที่กำหนดโดยเครื่องมือปัญญาประดิษฐ์ยังมีค่าไม่สูงมากนัก เนื่องจากโปรแกรมแต่ละตัวมีความสามารถที่โดดเด่นในภาพแต่ละหมวดหมู่แตกต่างกัน ในการกำหนดเมทาดาทาภาพถ่ายผู้ใช้งานจึงควรเลือกเครื่องมือให้เหมาะกับประเภทภาพถ่าย เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด นอกจากนี้จะพบว่าการใช้เครื่องมือปัญญาประดิษฐ์แต่เพียงอย่างเดียวไม่สามารถกำหนดเมทาดาทาที่ถูกต้องและสมบูรณ์ได้ มีความจำเป็นต้องใช้การตรวจสอบจากมนุษย์ในขั้นตอนสุดท้ายด้วย

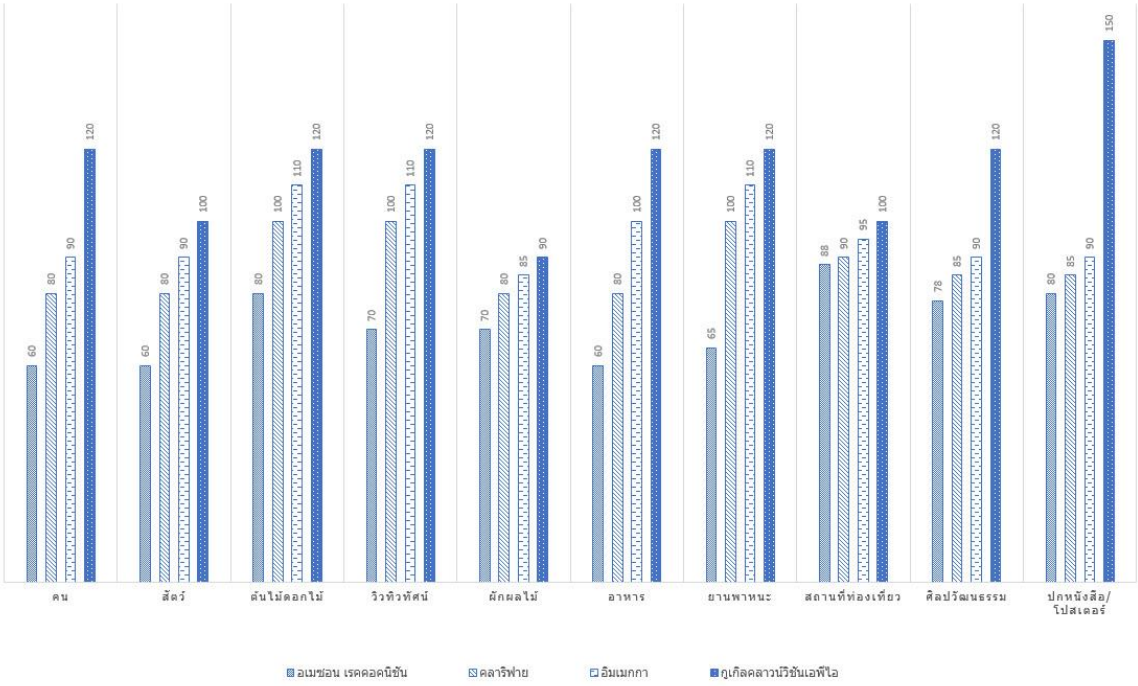
พิจารณาค่าเฉลี่ยของค่ามากสุดของภาพแต่ละประเภท จะได้ค่าเฉลี่ยค่าคำทับซ้อน ค่าเฉลี่ยคำที่ไม่ทับซ้อน ค่าเฉลี่ยค่าความคล้ายโคไซน์ และค่าเฉลี่ยความถูกต้องเอฟวันมีค่าเท่ากับ 75.59% 116% 76.84% และ 83.59% ตามลำดับ

ผลการทดลองเครื่องมือปัญญาประดิษฐ์ที่เหมาะสมกับภาพถ่ายแต่ละประเภท

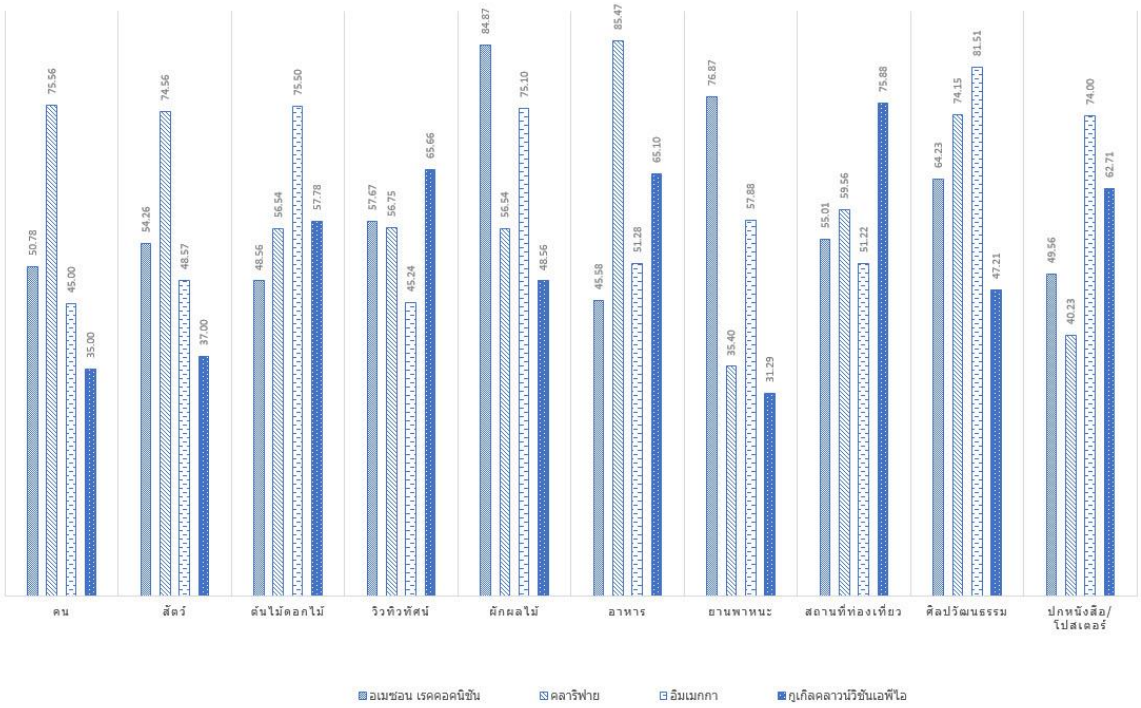
ผลการทดลองค่าคำทับซ้อน ค่าคำไม่ทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวันของภาพถ่ายแต่ละประเภทด้วยโปรแกรมอเมซอน เรคคอคนิชัน คลาริฟาย อิมเมกกา และกูเกิลคลาวน์วิชันเอพีไอแสดงในภาพที่ 32 ถึง ภาพที่ 35 ตามลำดับ



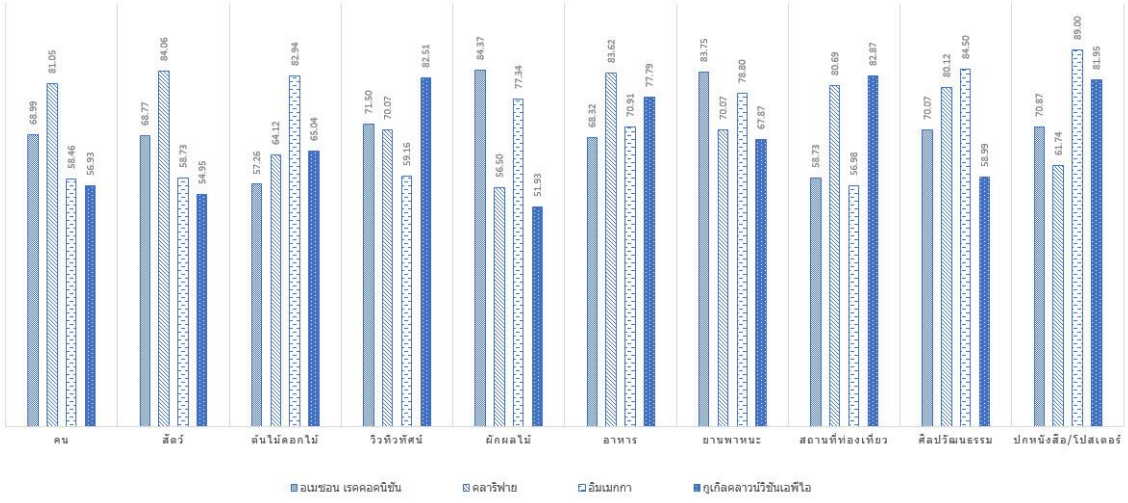
ภาพที่ 32 ค่าคำทับซ้อนของภาพแต่ละประเภท



ภาพที่ 33 ค่าคำที่ไม่ทับซ้อนของภาพแต่ละประเภท



ภาพที่ 34 การวัดค่าความคล้ายโคไซน์ของภาพแต่ละประเภท



ภาพที่ 35 ค่าความถูกต้องเอฟวันของภาพแต่ละประเภท

จากผลการทดลองทั้งหมดจะเห็นได้ว่าเครื่องมือแต่ละโปรแกรม สามารถกำหนดเมทาดาทาที่เกี่ยวข้องกับภาพถ่ายได้ โดยมีคุณสมบัติที่โดดเด่นแตกต่างกันในการกำหนดเมทาดาทาของภาพถ่ายแต่ละประเภท และผลการศึกษาพบว่าเมทาดาทาที่กำหนดโดยเครื่องมือปัญญาประดิษฐ์สามารถใช้ในการเพิ่มประสิทธิภาพการกำหนดเมทาดาทาภาพถ่ายได้สูงสุดถึง 150% ซึ่งสามารถนำมาใช้เพื่อเพิ่มประสิทธิภาพในการจัดหมวดหมู่ การทำดัชนี และการค้นหาภาพได้อีกด้วย โดยในการวิจัยนี้ได้สรุปโปรแกรมที่เหมาะสมกับรูปถ่ายแต่ละประเภทโดยพิจารณาจากค่าคำทับซ้อน การวัดค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวันที่สูงที่สุด สรุปได้ดังตาราง

ตารางที่ 69 สรุปเครื่องมือที่เหมาะสมกับภาพถ่ายแต่ละประเภท

หมวดหมู่ภาพ	**อเมซอน เรคคอคนิชัน**	**คลาริฟาย**	**อิมเมกกา**	**กูเกิลคลาวน์วิชันเอพีไอ**
คน		✓		
สัตว์		✓		
ต้นไม้ ดอกไม้			✓	
วิวทิวทัศน์				✓
ผัก ผลไม้	✓			
อาหาร		✓		
ยานพาหนะ	✓			
สถานที่ท่องเที่ยว				✓
ศิลปวัฒนธรรม			✓	
ปกหนังสือ/ โปสเตอร์เก่า			✓	

ตารางที่ 70 ลำดับเครื่องมือที่เหมาะสมกับภาพถ่ายแต่ละประเภท

หมวดหมู่ภาพ	อเมซอน เรคคอคนิชัน	คลาริฟาย	อิมเมกกา	กูเกิลคลาวน์วิชันเอพีไอ
คน	2	①	3	4
สัตว์	2	①	3	4
ต้นไม้ ดอกไม้	4	3	①	2
วิวทิวทัศน์	2	3	4	①
ผัก ผลไม้	①	3	2	4
อาหาร	4	①	3	2
ยานพาหนะ	①	3	2	4
สถานที่ท่องเที่ยว	3	2	4	①
ศิลปวัฒนธรรม	3	2	①	4
ปกหนังสือ/ โปสเตอร์เก่า	3	4	①	2

ตารางที่ 71 โปรแกรมที่ให้ค่าความถูกต้องสูงสุดของแต่ละหมวดภาพ

โปรแกรม	หมวดภาพ
อเมซอน เรคคอคนิชัน	ผัก ผลไม้ / ยานพาหนะ
คลาริฟาย	คน / สัตว์ / อาหาร
อิมเมกกา	ต้นไม้ ดอกไม้ / ศิลปวัฒนธรรม / ปกหนังสือ โปสเตอร์เก่า
กูเกิลคลาวน์วิชันเอพีไอ	วิวทิวทัศน์ / สถานที่ท่องเที่ยว

ตารางที่ 70 แสดงลำดับเครื่องมือที่เหมาะกับภาพถ่ายแต่ละประเภท เรียงลำดับจากเหมาะสมมากที่สุดไปน้อยที่สุด ด้วยเลข 1 ถึง 4 และตารางที่ 71 แสดงโปรแกรมที่ให้ค่าความถูกต้องสูงสุดของแต่ละหมวดภาพ

บทที่ 5
สรุป อภิปรายผล และข้อเสนอแนะ

สรุป

จากการเปลี่ยนแปลงทางเทคโนโลยีที่เกิดขึ้นในปัจจุบัน ข้อมูลภาพถ่ายถูกบันทึกไว้เป็นจำนวนมาก เป็นชุดข้อมูลที่ใหญ่ขึ้นเรื่อย ๆ จากการที่มีเครื่องมือที่ใช้บันทึกภาพได้ในราคาถูก พกพาง่าย และมีประสิทธิภาพเทียบเท่ากับกล้องดิจิทัลราคาแพง การจัดเก็บข้อมูลภาพถ่ายเหล่านี้เพื่อใช้งานในอนาคตจึงเป็นสิ่งที่จำเป็นต้องพิจารณาถึงวิธีการที่เหมาะสม การจัดเก็บภาพถ่ายเพื่อการเข้าถึงได้อย่างรวดเร็วนั้นจำเป็นต้องมีการเก็บข้อมูล ลักษณะสำคัญของภาพถ่ายไว้ด้วย ซึ่งข้อมูลนี้เรียกอีกอย่างหนึ่งว่าเมทาดาทา การจัดเตรียมข้อมูลเมทาดาทาคุณภาพสูง หรือกล่าวอีกอย่างหนึ่งว่าเมทาดาทาที่สมบูรณ์ เป็นสิ่งสำคัญสำหรับการจัดระเบียบการเข้าถึงภาพถ่าย เพื่อให้สามารถเข้าถึงภาพถ่ายที่มีส่วนประกอบของวัตถุที่แตกต่างกัน เมทาดาทาที่กำหนดจึงต้องสามารถระบุวัตถุที่หลากหลายในภาพได้ การใช้เครื่องมือปัญญาประดิษฐ์เพื่อกำหนดเมทาดาทารูปภาพเป็นบริการที่แพร่หลาย และได้รับความนิยมเพิ่มมากขึ้น การตรวจจับวัตถุของแต่ละเครื่องมือมีประสิทธิภาพที่แตกต่างกันขึ้นอยู่กับจำนวนข้อมูลที่ใช้ในการฝึกอัลกอริทึมที่ใช้ในการฝึกสอน และทรัพยากรทางระบบคอมพิวเตอร์ต่าง ๆ

งานวิจัยนี้นำเสนอการศึกษาการใช้ปัญญาประดิษฐ์เพื่อพัฒนาเมทาดาทาภาพถ่ายดิจิทัล โดยการเปรียบเทียบประสิทธิภาพของเมทาดาทาที่ได้จากโปรแกรมปัญญาประดิษฐ์ 4 โปรแกรม ได้แก่ อเมซอน เรคคอคนิชัน (Amazon Rekognition) คลาริฟาย (Clarifai) อิมเมกกา (Imagga) และกูเกิลคลาวน์วิชัน เอพีไอ (Google cloud vision API) โดยทดลองกับภาพถ่ายที่เก็บรวบรวมจากคลังภาพดิจิทัลแบบเปิดเผยแพร่ภายใต้โดเมนสาธารณะ (public domain) ซึ่งไม่มีลิขสิทธิ์ หรือเจ้าของลิขสิทธิ์ยกให้เป็นสาธารณะประโยชน์เวปไซต์ที่เผยแพร่ข้อมูลจำนวน 1,000 ภาพแบ่งเป็น 10 หมวดหมู่ๆ หมวดหมู่ละ 100 ภาพ ประกอบด้วยภาพคน สัตว์ ต้นไม้ดอกไม้ วิวทิวทัศน์ ผักผลไม้ อาหาร ยานพาหนะ สถานที่ท่องเที่ยว และศิลปวัฒนธรรม การทดลองใช้การเปรียบเทียบเมทาดาทาที่ได้มาพร้อมภาพถ่าย กับเมทาดาทาที่ได้จากเครื่องมือ โดยทำการวิเคราะห์แยกตามหมวดหมู่ภาพแต่ละประเภท ความถูกต้องและความสมบูรณ์ของเมทาดาทาที่สร้างโดยเครื่องมือเหล่านี้สามารถเปรียบเทียบ วิเคราะห์และประเมินประสิทธิภาพด้วยค่าทางสถิติต่าง ๆ ได้แก่ การวัดค่าคำทับซ้อน (Word Overlap Similarity Measure) การวัดค่าคำที่ไม่ทับซ้อน (No Word Overlap Similarity Measure) การวัดค่าความคล้ายโคไซน์ (Cosine Similarity) การหาค่าความไว (Recall) ค่าความแม่นยำ (Precision) และ ค่าความถูกต้องเอฟวัน (F1-score) จากการศึกษาสรุปได้ดังนี้

1. ค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวันที่มากที่สุดของทุกหมวดหมู่ภาพมีค่าเฉลี่ยเท่ากับ 83.59% ซึ่งมากกว่าเกณฑ์ที่ 80% แสดงให้เห็นว่าเครื่องมือปัญญาประดิษฐ์สามารถนำมาใช้กำหนดเมทาดาทาจากภาพถ่ายเพื่อเพิ่มประสิทธิภาพเมทาดาทาได้

2. โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ กำหนดเมทาดาทาเป็นคำที่ไม่ทับซ้อนแต่มีความหมายที่เกี่ยวข้องเหมาะสม ช่วยเพิ่มโอกาสในการค้นหาได้สูงสุดในทุกหมวดภาพ มีค่าเฉลี่ยที่ 116% และมีค่าสูงสุดถึง 150% ในหมวดปกหนังสือ/โปสเตอร์เก่า

3. เครื่องมือที่เหมาะสมในการจัดการกับภาพถ่ายแต่ละประเภทโดยพิจารณาค่าคำทับซ้อน ค่าความคล้ายโคไซน์ และ ค่าความถูกต้องเอฟวันที่สูงที่สุด แยกตามหมวดหมู่ภาพดังนี้

อเมซอน เรคคอคนิชัน (Amazon Rekognition) เหมาะกับการใช้กำหนดเมทาดาทาภาพถ่ายผักผลไม้ และยานพาหนะ

คลาริฟาย (Clarifai) เหมาะกับการใช้กำหนดเมทาดาทาภาพถ่ายคน สัตว์และอาหาร

อิมเมกกา (Imagga) เหมาะกับการใช้กำหนดเมทาดาทาภาพถ่ายต้นไม้ดอกไม้ ศิลปวัฒนธรรม และปกหนังสือ/โปสเตอร์เก่า

กูเกิลคลาวน์วิชันเอพีไอ (Google cloud vision API) เหมาะกับการใช้กำหนดเมทาดาทาภาพถ่ายวิวทิวทัศน์ และสถานที่ท่องเที่ยว

เครื่องมือปัญญาประดิษฐ์สามารถนำมาใช้ในการกำหนดเมทาดาทาภาพถ่ายร่วมกับเมทาดาทาที่กำหนดโดยมนุษย์ได้อย่างมีประสิทธิภาพ เครื่องมือแต่ละชนิดมีชุดคุณสมบัติ จุดแข็งและจุดอ่อนของตัวเอง การวิจัยนี้ได้นำเสนอข้อมูลเกี่ยวกับประสิทธิภาพของการกำหนดเมทาดาทาจากเครื่องมือปัญญาประดิษฐ์ โดยผลลัพธ์ชี้ให้เห็นว่าเครื่องมือทั้งสี่สามารถสร้างข้อมูลเมทาดาทาที่ถูกต้องและสมบูรณ์ได้ โดยมีผลลัพธ์ที่แตกต่างกันในภาพถ่ายแต่ละประเภท

โดยสรุปเครื่องมือกำหนดเมทาดาทาที่ดีที่สุดจะขึ้นอยู่กับความต้องการเฉพาะของผู้ใช้ และการยอมรับค่าความแม่นยำและความไว ในบางกรณีความเที่ยงตรงสูงอาจมีความสำคัญมากกว่า ในขณะที่บางกรณีความไวที่สูงอาจเป็นที่ต้องการมากกว่า ในกรณีของงานวิจัยนี้ พิจารณาเลือกเครื่องมือที่ดีที่สุดสำหรับภาพแต่ละประเภทจากค่าคำทับซ้อน การวัดค่าความคล้ายโคไซน์ และค่าความถูกต้องเอฟวัน ซึ่งเป็นค่าเฉลี่ยของความแม่นยำและความไว เนื่องจากมีความสมดุลระหว่างความแม่นยำและความไว การเลือกเครื่องมือสำหรับการใช้งานนอกจากผลลัพธ์เชิงปริมาณเหล่านี้แล้ว สิ่งสำคัญคือต้องพิจารณาปัจจัยอื่น ๆ ร่วมด้วยเช่น ปริมาณภาพถ่าย ต้นทุน เวลาในการประมวลผล คุณภาพและความละเอียดของภาพ ความสะดวกในการใช้งาน ระดับความเชี่ยวชาญทางเทคนิคที่จำเป็นสำหรับแต่ละเครื่องมือ ความคิดเห็นของผู้ใช้ และประสบการณ์ควรนำมาพิจารณาด้วยเมื่อตัดสินใจขั้นสุดท้ายว่าจะใช้เครื่องมือใด

การกำหนดเมทาดาทาจากมนุษย์เองนั้นมีทั้งข้อดี และข้อจำกัด ข้อดีคือมนุษย์สามารถระบุเมทาดาทาที่เฉพาะเจาะจง เช่นชื่อสถานที่ ชื่อบุคคล สถานการณ์ ชื่อเรียกเฉพาะ ได้อย่างถูกต้อง แต่มีข้อจำกัดในเรื่องของการใช้คำที่ไม่หลากหลาย ไม่ครอบคลุมความหมายทั้งหมดที่อาจถูกตีความได้แตกต่างกันในผู้ใช้แต่ละคน ในขณะที่เทคโนโลโยปัญญาประดิษฐ์มีข้อจำกัดด้านความสามารถในการกำหนดเมทาดาทาเฉพาะได้ไม่สมบูรณ์ แต่สามารถช่วยในการวิเคราะห์คำที่มีความหมายคล้ายคลึงกันจำนวนมาก และถูกต้องให้ได้ อีกทั้งการใช้เครื่องมือปัญญาประดิษฐ์จะช่วยให้สามารถดึงข้อมูลเชิงลึก หรือข้อมูลที่ไม่เป็นที่รู้จักในปัจจุบันจากภาพถ่ายได้ ช่วยเพิ่มประสิทธิภาพคำค้น เพื่อให้สืบค้นภาพถ่ายในถูกต้องมากขึ้น ดังนั้นการทำงานร่วมกันของมนุษย์และปัญญาประดิษฐ์จึงยังเป็นความจำเป็นในการทำงานในอนาคต

อภิปรายผล

การศึกษาการใช้เครื่องมือปัญญาประดิษฐ์ในการกำหนดเมทาดาทาภาพถ่ายนี้ พบว่าการใช้เครื่องมือปัญญาประดิษฐ์สามารถเพิ่มประสิทธิภาพของเมทาดาทา และสามารถนำเมทาดาทาที่ได้มาช่วยในการจัดประเภทข้อมูลภาพถ่ายได้ สอดคล้องกับงานวิจัยของเบกเกอร์ (Bakker et al., 2020) บินมาฆะเซิน (Binmakhashen & Mahmoud, 2020) กาซาลี (Ghazali et al., 2020) และโคลาวิซซ่า (Colavizza et al, 2021) การใช้เครื่องมือทางปัญญาประดิษฐ์ในการกำหนดเมทาดาทาจะสามารถช่วยในการจัดระเบียบ ค้นหา และเข้าถึงข้อมูลภาพถ่ายได้อย่างมีประสิทธิภาพ และรวดเร็ว อีกทั้งยังสามารถช่วยในการการวิเคราะห์ การตรวจสอบภาพถ่าย การจัดทำดัชนี คอลเลกชันดิจิทัล และหมวดหมู่ภาพถ่ายแบบอัตโนมัติจากการตรวจจับวัตถุต่างๆ ในภาพได้อีกด้วย การนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ในระบบจัดการข้อมูลภาพถ่ายจดหมายเหตุเป็นงานที่มีความสำคัญ และจำเป็นต้องมีการพัฒนาต่อไปในอนาคต เมทาดาทาที่ถูกต้อง และครบถ้วนจะช่วยให้การตรวจสอบถูกต้องของภาพดิจิทัลทำได้อย่างมีประสิทธิภาพมากยิ่งขึ้น สอดคล้องกับงานวิจัยของฮาร์ราน (Harran et al., 2017) และ กังวาล (Gangwar & Pathania, 2018) บึทเนอร์ (Büttner, 2019) และสุชมา (Sushma & Kumari, 2021) และผลการทดลองสอดคล้องกับคิเรียโค (Kyriakou et al., 2019) ที่นำเสนอว่าโปรแกรมคลาริฟายกำหนดเมทาดาทาของมนุษย์ได้แม่นยำที่สุดเมื่อเปรียบเทียบกับโปรแกรมอิมเมกกา และ กูเกิลคลาวน์วิชันเอพีไอ

หากพิจารณาถึงอัลกอรึทึม หรือขั้นตอนวิธีที่แต่ละโปรแกรมใช้ ซึ่งเป็นปัจจัยหนึ่งของความสำเร็จในการกำหนดเมทาดาทา พบว่าทั้ง 4 โปรแกรมได้แก่อเมซอน เรคคอคนิชัน คลาริฟาย อิมเมกกา และ กูเกิลคลาวน์วิชันเอพีไอ ใช้เทคนิคการเรียนรู้เชิงลึกทั้งหมด แต่เนื่องจากว่าเป็นโปรแกรมเชิงพาณิชย์จึงไม่ได้เปิดเผยถึงรายละเอียดของเทคนิคที่ใช้ ถึงแม้จะใช้อัลกอรึทึมการเรียนรู้เชิงลึกเหมือนกันแต่ได้ค่าความแม่นยำที่แตกต่างกัน เกิดได้จากหลายสาเหตุ ดังนี้

1. คุณภาพและขนาดของชุดข้อมูลที่ใช้สอนอัลกอริทึมมีผลกระทบอย่างมากต่ออัตราความแม่นยำ ชุดข้อมูลที่ใหญ่และหลากหลายมากขึ้นจะสามารถช่วยให้อัลกอริทึมเรียนรู้ลักษณะและรายละเอียดต่างๆ มากขึ้น ซึ่งสามารถทำให้มีความแม่นยำสูงขึ้นได้ ในทางตรงกันข้าม

ชุดข้อมูลที่เล็กหรือมีความเอียงมากอาจจำกัดความสามารถของอัลกอริทึมในการสรุปผลเพื่อใช้กับข้อมูลใหม่ๆ ซึ่งอาจทำให้อัตราความแม่นยำลดลงได้

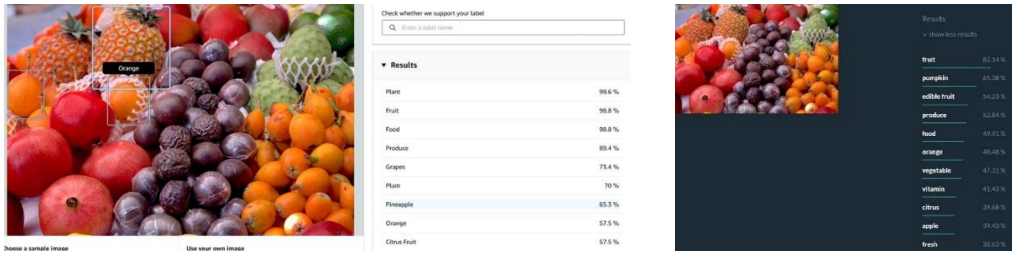
2. การเลือกค่าพารามิเตอร์หรือฮายเปอร์พารามิเตอร์ เช่น อัตราการเรียนรู้ (learning rate), ขนาดชุดข้อมูล (batch size), และจำนวนชั้น (number of layers) ยังสามารถมีผลต่ออัตราความแม่นยำของอัลกอริทึมได้ พารามิเตอร์เหล่านี้เป็นพารามิเตอร์ที่สามารถปรับได้โดยไม่ต้องเรียนรู้ในระหว่างกระบวนการสอน แต่จะต้องถูกกำหนดโดยนักพัฒนาโปรแกรม การเลือกค่าพารามิเตอร์ที่เหมาะสมสามารถช่วยเพิ่มอัตราความแม่นยำได้ ในขณะที่การเลือกค่าพารามิเตอร์ที่ไม่เหมาะสมอาจส่งผลต่ออัตราความแม่นยำได้
3. ความแตกต่างในฮาร์ดแวร์หรือซอฟต์แวร์ที่ใช้ในการฝึกอัลกอริทึมก็สามารถมีผลต่ออัตราความแม่นยำของมันได้ด้วย ฮาร์ดแวร์ที่แตกต่างกัน เช่น CPUs, GPUs, หรือ TPUs อาจมีความเร็วในการประมวลผลและความจำที่แตกต่างกัน ซึ่งสามารถมีผลต่อกระบวนการฝึกอบรมได้ นอกจากนี้ ไลบรารีหรือเวอร์ชันซอฟต์แวร์ที่แตกต่างกันอาจทำให้มีความแม่นยำที่แตกต่างกันได้

จากการศึกษาได้ทำการอภิปรายผลลัพธ์แยกแต่ละโปรแกรมดังนี้

อเมซอน เรคคอคนิชัน

อเมซอน เรคคอคนิชัน เป็นเครื่องมือที่มีความสามารถหลากหลาย จุดแข็งหลักประการหนึ่งของคือความสามารถในการระบุวัตถุ ฉาก และกิจกรรมภายในรูปภาพและวิดีโอได้อย่างแม่นยำ โดยใช้อัลกอริธึมการเรียนรู้เชิงลึกที่ได้รับการฝึกฝนในชุดข้อมูลรูปภาพและวิดีโอขนาดใหญ่ จุดแข็งอีกประการคือความสามารถในการวิเคราะห์ใบหน้า รวมถึงการจดจำใบหน้าและการตรวจจับลักษณะใบหน้า ซึ่งช่วยให้เครื่องมือสามารถระบุและดึงข้อมูลเกี่ยวกับบุคคลภายในรูปภาพและวิดีโอ เช่น เพศ อายุ และอารมณ์ ซึ่งมีประโยชน์ในการใช้งานที่หลากหลาย เช่น การรักษาความปลอดภัยและการเฝ้าระวัง การตลาดและการโฆษณา และการบริการลูกค้า นอกจากความสามารถในการวิเคราะห์รูปภาพและวิดีโอแล้วยังมีความสามารถในการวิเคราะห์ข้อความอีกด้วย ซึ่งช่วยให้เครื่องมือสามารถแยกข้อความจากรูปภาพและวิดีโอได้โดยอัตโนมัติ และทำการวิเคราะห์ความคิดเห็นต่อข้อความที่แยกออกมา ซึ่งมีประโยชน์ในแอปพลิเคชันต่างๆ เช่น การตรวจสอบโซเชียลมีเดียและการวิเคราะห์ความคิดเห็นของลูกค้า นอกจากนี้ยังสามารถรวมบริการเข้ากับบริการอื่นๆ ของอเมซอน เช่น อเมซอนเอสทรี (Amazon S3) และ อเมซอนแลมดา (Amazon Lambda) ได้ง่าย ช่วยให้ปรับใช้ได้อย่างราบรื่นและยืดหยุ่น โดยรวมแล้วอเมซอนเรคคอคนิชัน เป็นเครื่องมือที่มีประสิทธิภาพสำหรับการสร้างข้อมูลเมทาดาทาที่มีความสามารถหลากหลาย ความสามารถในการระบุวัตถุ ฉาก และกิจกรรมภายในภาพและวิดีโอได้อย่างแม่นยำ ทำการวิเคราะห์ใบหน้า และดึงข้อความจากภาพและวิดีโอ เมื่อเปรียบเทียบผลการทดลองกับโปรแกรมอื่น ในตัวอย่างภาพผลไม้พบว่าถึงแม้ว่าโปรแกรมอเมซอน เรคคอคนิชัน จะกำหนดเมทาดาทาที่ความเชื่อมั่นสูงสุดเป็น plant ไม่ใช่ vegetable แต่เมทาดาทาที่กำหนดได้ มีความถูกต้องของชนิดผลไม้ที่ตรวจพบ

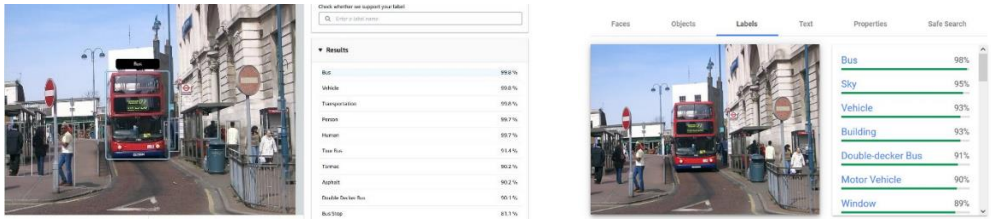
มากที่สุด ในขณะที่อิมเมกกา กำหนดเมทาดาทาที่ความเชื่อมั่นสูงสุดเป็น vegetable แต่กำหนดเมทาดาทาชนิดของผลไม้ในภาพผิด เช่น pumpkin



(ก) (ข)

ภาพที่ 36 เปรียบเทียบเมทาดาทาภาพผลไม้จากโปรแกรม (ก)อเมซอน เรคคอคนิชัน และ (ข)อิมเมกกา

จากตัวอย่างภาพยานพาหนะทั้งโปรแกรมอเมซอน เรคคอคนิชัน และกูเกิลคลาวน์วิชันเอพีไอ สามารถกำหนดเมทาดาทาระดับความเชื่อมั่นสูงสุดเป็น Bus ทั้งสองโปรแกรม แต่เมื่อดูเมทาดาทาลำดับต่อๆ มาจะพบว่าเมทาดาทาของโปรแกรมอเมซอน เรคคอคนิชันมีความถูกต้องและสัมพันธกับภาพมากกว่า



(ก) (ข)

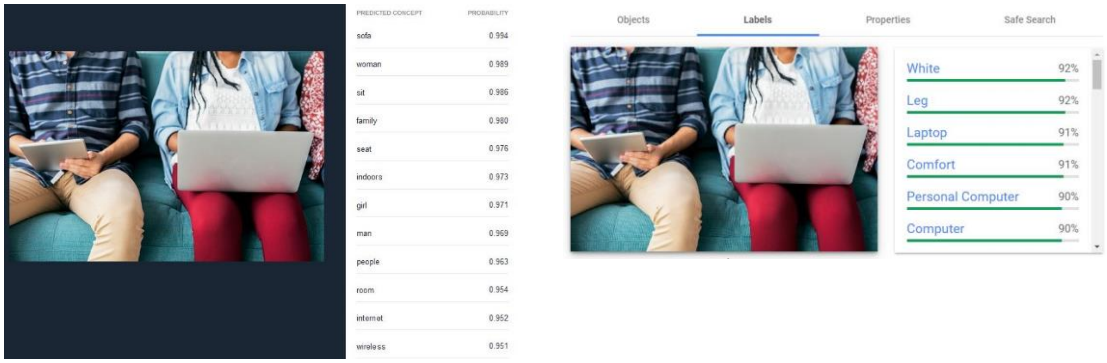
ภาพที่ 37 เปรียบเทียบเมทาดาทาภาพยานพาหนะจากโปรแกรม (ก) อเมซอน เรคคอคนิชัน และ (ข) กูเกิลคลาวน์วิชันเอพีไอ

คลาริฟาย

คลาริฟายมีประสิทธิภาพดีทั้งในด้านความแม่นยำและความไว จุดแข็งคือความสามารถในการจดจำเนื้อหารูปภาพที่หลากหลาย รวมถึงวัตถุ ฉาก และกิจกรรมต่างๆ ทำให้เหมาะสำหรับแอปพลิเคชันที่ต้องการระบุข้อมูลเมทาดาทาหลายประเภท เช่น การแท็กรูปภาพ การกรองเนื้อหา และการค้นหารูปภาพ นอกจากนี้เอพีไอของคลาริฟายยังใช้งานง่ายและรวมเข้ากับแอปพลิเคชันอื่นๆ ได้ง่าย ซึ่งทำให้เป็นตัวเลือกที่น่าสนใจสำหรับนักพัฒนาและธุรกิจที่ต้องการเพิ่มความสามารถในการจดจำรูปภาพให้กับผลิตภัณฑ์ของตัวเอง จุดแข็งอีกประการคือความสามารถในการเรียนรู้และปรับปรุงเมื่อเวลาผ่านไป ด้วยโมเดลแมชชีนเลิร์นนิงที่ได้รับการฝึกอบรมในชุดข้อมูลขนาดใหญ่ของรูปภาพ และสามารถปรับให้เข้ากับรูปภาพและบริบทใหม่ได้ ซึ่งหมายความว่าความแม่นยำจะดีขึ้นเมื่อมีการป้อนข้อมูลเข้าสู่ระบบมากขึ้น

ในทางกลับกัน ประสิทธิภาพของโมเดลจะขึ้นอยู่กับคุณภาพของชุดข้อมูลที่ได้รับการฝึกฝน ชุดข้อมูลที่เอนเอียงหรือไม่หลากหลายเพียงพออาจนำไปสู่ผลลัพธ์การจดจำที่ไม่ดี นอกจากนี้โมเดลอาจทำงานได้ไม่ดีเมื่อใช้กับโดเมนหรือบริบทที่ไม่ได้รับการฝึกอบรม โดยสรุปคลาริฟายเป็นเครื่องมือการจดจำรูปภาพที่มีประสิทธิภาพและหลากหลาย รองรับเนื้อหาภาพที่ที่หลากหลายและ มีส่วนการใช้งานที่เป็นมิตรกับผู้ใช้

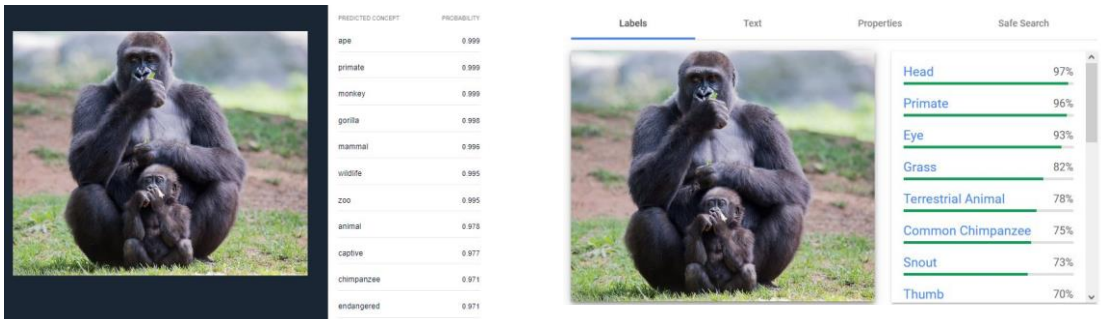
เมื่อเปรียบเทียบผลการทดลองกับโปรแกรมอื่นในตัวอย่างภาพคนโปรแกรมคลาริฟาย สามารถกำหนดเมทาดาทาระบุคน “man” “girl” “people” ได้รวมถึงลักษณะการนั่ง “sit” ในขณะที่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอนำเสนอเมทาดาทาไปที่สิ่งของมากกว่าเช่น “computer” “laptop” leg”



(ก) (ข)

ภาพที่ 38 เปรียบเทียบเมทาดาทาภาพคนจากโปรแกรม (ก)คลาริฟาย และ (ข)กูเกิลคลาวน์วิชันเอพีไอ

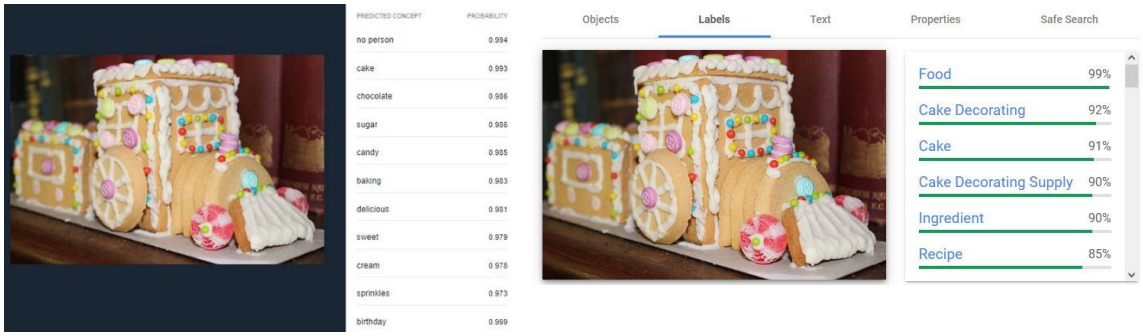
ในตัวอย่างภาพสัตว์ถึงแม้ว่าโปรแกรมคลาริฟาย สามารถกำหนดเมทาดาทาระบุเผ่าพันธ์ได้ “ape” “primate” “gorilla” ในขณะที่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอนำเสนอเมทาดาทาไปที่ส่วนประกอบมากกว่าเช่น “head” “eye” “grass”



(ก) (ข)

ภาพที่ 39 เปรียบเทียบเมทาดาทาภาพสัตว์จากโปรแกรม (ก)คลาริฟาย และ (ข)กูเกิลคลาวน์วิชันเอพีไอ

ในตัวอย่างภาพอาหารถึงแม้ว่าโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ จะกำหนดเมทาดาทา Food ที่ระดับความเชื่อมั่นสูงสุด ในขณะที่คลาริฟายสามารถกำหนดเมทาดาทา cake ได้ในระดับความเชื่อมั่นลำดับสอง แต่คลาริฟายสามารถกำหนดเมทาดาทาที่เป็นรายละเอียดของเค้กได้ละเอียดมากที่สุด เช่น candy cream sprinkles



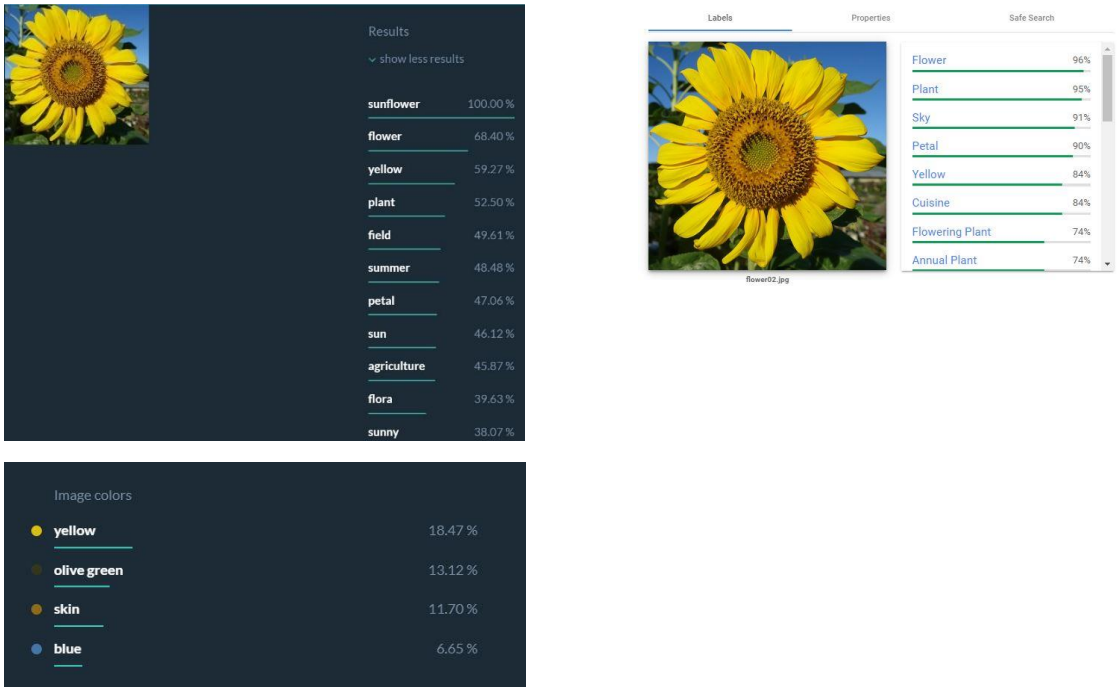
(ก) (ข)

ภาพที่ 40 เปรียบเทียบเมทาดาทาภาพอาหารจากโปรแกรม (ก)คลาริฟาย และ (ข)กูเกิลคลาวน์วิชันเอพีไอ

อิมเมกกา

อิมเมกกา เป็นเครื่องมือจดจำและวิเคราะห์รูปภาพที่มีประสิทธิภาพ จุดแข็งคือความสามารถในการแท็กรูปภาพโดยอัตโนมัติด้วยข้อมูลเมทาดาทาที่เกี่ยวข้อง เช่น ป้ายวัตถุและฉาก ข้อมูลสี และอื่น ๆ ช่วยให้ผู้ใช้สามารถค้นหาและจัดระเบียบไลบรารีรูปภาพของตนได้อย่างสะดวกและรวดเร็ว มีประโยชน์สำหรับอุตสาหกรรมต่าง ๆ เช่น อีคอมเมิร์ซ สื่อและความบันเทิง และอสังหาริมทรัพย์ จุดแข็งอีกประการคือความสามารถในการจดจำและวิเคราะห์ภาพตามขนาด ด้วยการรองรับการประมวลผลภาพจำนวนมากพร้อมกัน ทำให้สามารถวิเคราะห์ภาพนับพันหรือล้านภาพได้อย่างรวดเร็วและมีประสิทธิภาพ เหมาะสำหรับองค์กรที่มีคลังภาพขนาดใหญ่ นอกเหนือจากการติดแท็กรูปภาพและการจัดระเบียบแล้ว อิมเมกกายังมีความสามารถในการวิเคราะห์รูปภาพที่หลากหลาย รวมถึงการครอบตัดรูปภาพ การปรับขนาดรูปภาพ และการปรับปรุงรูปภาพ

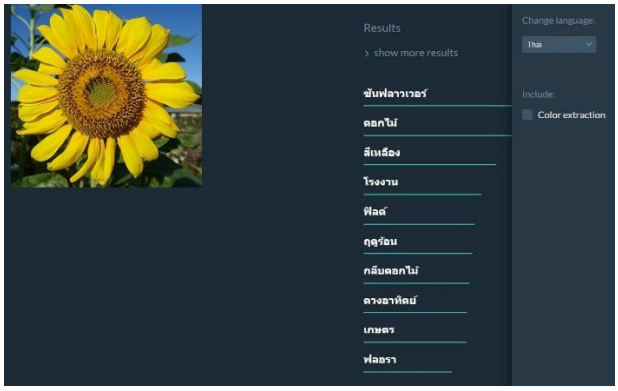
จากผลเมทาดาทาตัวอย่างภาพดอกไม้ โปรแกรมอิมเมกกา สามารถกำหนดประเภทของดอกไม้ sunflower ในระดับความเชื่อมั่นสูงสุดได้ถูกต้อง ในขณะที่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ กำหนดคำที่ระดับความเชื่อมั่นสูงเป็นส่วนประกอบแวดล้อม ไม่ใช่ส่วนประกอบหลัก เช่นคำว่า sunflower กำหนดได้แต่เป็นที่ระดับความเชื่อมั่นที่ต่ำลงมา



(ก) (ข)

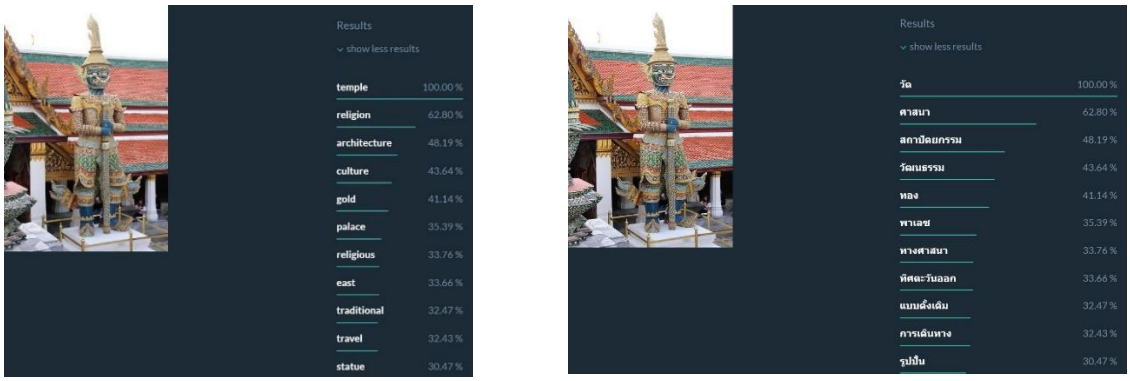
ภาพที่ 41 เปรียบเทียบเมทาดาทาภาพดอกไม้จากโปรแกรม (ก)อิมเมกกา และ (ข)กูเกิลคลาวน์วิชันเอพีไอ

ในเรื่องความหลากหลายของภาษา โปรแกรมอื่นสามารถกำหนดเมทาดาดาได้หลายภาษาเหมือนกัน แต่มีเพียงโปรแกรมอิมเมกกาที่มีการแสดงผลเมทาดาทาภาษาไทยได้ชัดเจน ถึงแม้ผลภาษาไทยที่ได้มาจะยังไม่ถูกต้องมากนัก ดังภาพที่ 41 ในภาพดอกไม้ใช้การทับศัพท์ “ซันฟลาวเวอร์” แทนการใช้คำว่า “ดอกทานตะวัน” หรือในภาพยักษ์ที่ใช้คำทับศัพท์ “พาเลซ” แทนการใช้ว่า “พระราชวัง”

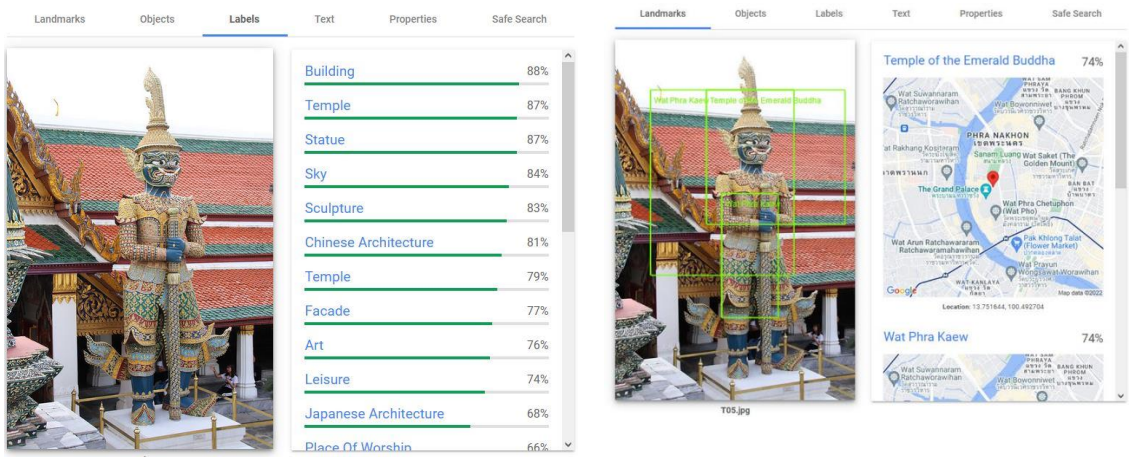


ภาพที่ 42 ผลการกำหนดเมทาดาทาภาษาไทยของภาพดอกไม้จากโปรแกรมอิมเมกกา

จากผลเมทาดาทาตัวอย่างศิลปวัฒนธรรม โปรแกรมอิมเมกกาสามารถกำหนดประเภทสถานที่ “temple” ได้ในระดับความเชื่อมั่นสูงสุดได้ถูกต้อง อีกทั้งยังสามารถระบุรายละเอียดของชิ้นงานเช่น “palace” “gold” “status” ได้อย่างถูกต้อง ในขณะที่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ หากพิจารณาแค่เมทาดาทาที่กำหนดมา คำที่ระดับความเชื่อมั่นสูงเป็นคำที่แสดงส่วนประกอบแวดล้อมไม่ใช่ส่วนหลัก แต่หากนำความสามารถด้านการกำหนดตำแหน่งที่ตั้ง (Landmark) ที่ได้คือ “Temple of the Emerald Buddha” มาร่วมด้วยจะทำให้ได้เมทาดาทาที่มีความสมบูรณ์มากขึ้น



ภาพที่ 43 ผลการกำหนดเมทาดาทาภาพศิลปวัฒนธรรม ทั้งภาษาอังกฤษ และภาษาไทยจากโปรแกรมอิมเมกกา



ภาพที่ 44 ผลเมทาดาทา และตำแหน่งที่ตั้ง ภาพศิลปวัฒนธรรมจากโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ

จากผลเมทาดาทาตัวอย่างภาพโปสเตอร์เก่า โปรแกรมอิมเมกกาสามารถระบุคำที่ใกล้เคียงวัตถุในภาพเช่น “covering” “mask” “protective covering” ได้ในระดับความเชื่อมั่นสูง ในขณะที่โปรแกรมกูเกิลคลาวน์วิชันเอพีไอกำหนดเมทาดาทาระดับความเชื่อมั่นสูงเป็นส่วนประกอบของใบหน้า เช่น “nose” “jaw” “wrist” “beard” แต่คำสำคัญอย่าง “gas mask” ได้ในระดับความเชื่อมั่นต่ำ

(ก) (ข)

ภาพที่ 45 เปรียบเทียบเมทาดาทาภาพโปสเตอร์จากโปรแกรม (ก)อิมเมกกา และ (ข)กูเกิลคลาวน์วิชันเอพีไอ

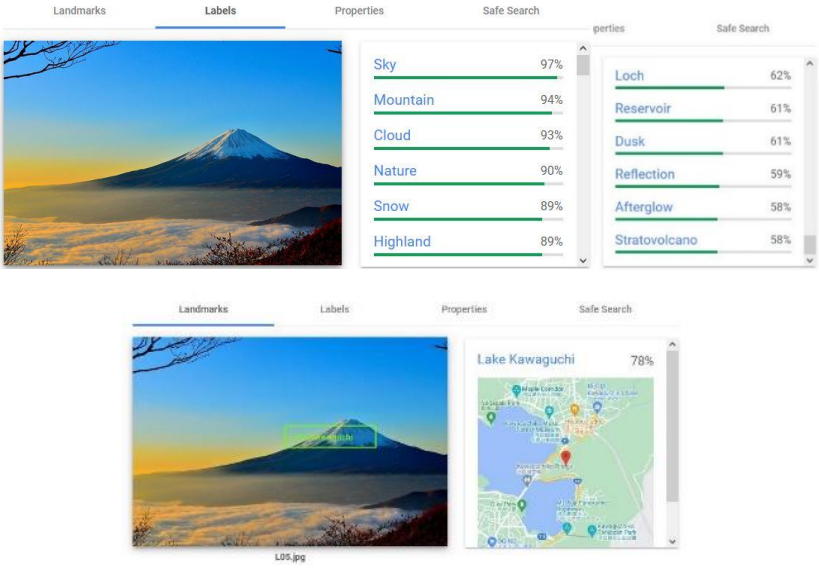
กูเกิลคลาวน์วิชันเอพีไอ

กูเกิลคลาวน์วิชันเอพีไอ เป็นบริการจดจำรูปภาพบนคลาวด์ที่ให้บริการโดยกูเกิล จุดแข็งที่สำคัญคือความสามารถในการทำงานการจดจำรูปภาพที่หลากหลาย รวมถึงการตรวจจับวัตถุ การจดจำใบหน้าและการดึงข้อความ สามารถใช้เป็นเครื่องมือในการรักษาความปลอดภัย ไปจนถึงการค้นหาด้วยภาพและการทำอีคอมเมิร์ซ นอกจากนี้ยังช่วยให้นักพัฒนาสามารถรวมฟังก์ชันเข้ากับแอปพลิเคชันได้อย่างง่ายเป็นตัวเลือกที่สะดวกสำหรับธุรกิจและองค์กรที่ต้องการเพิ่มความสามารถในการจดจำรูปภาพให้กับผลิตภัณฑ์และบริการของตน จุดแข็งอีกอย่างคือความสามารถในการปรับขนาด ในฐานะที่เป็นบริการบนคลาวด์สามารถจัดการข้อมูลและการให้บริการจำนวนมากได้ ทำให้เหมาะสำหรับการใช้งานในแอพพลิเคชั่นที่มีปริมาณมาก ปัญหาหนึ่งที่อาจเกิดขึ้นคือค่าใช้จ่าย เนื่องจากการใช้เอพีไอมีค่าใช้จ่ายสูงหากมีการใช้งานสูง นอกจากนี้ประสิทธิภาพของเอพีไออาจแตกต่างกันไปขึ้นอยู่กับคุณภาพและความละเอียดของภาพที่ใช้และอาจไม่ได้ผลลัพธ์ที่แม่นยำเสมอไปในกรณีที่คุณภาพของภาพไม่ดี เช่น ภาพที่มีความละเอียดต่ำ หรือมีสัญญาณรบกวน การเลือกใช้งานสิ่งสำคัญที่ต้องพิจารณาประกอบด้วยคือต้นทุนและประสิทธิภาพ และประเมินกับทางเลือกอื่น ๆ เพื่อประกอบการตัดสินใจ

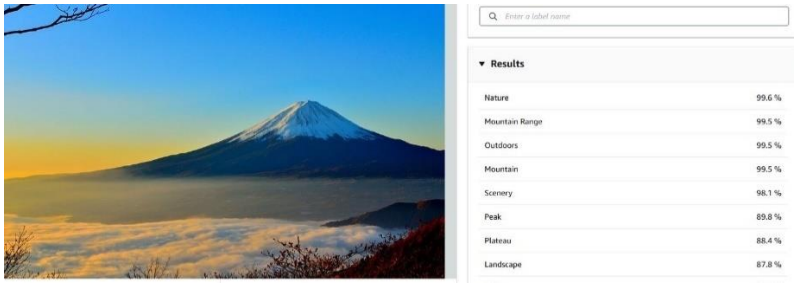
จากการทดลองพบว่าเมทาดาทาที่กำหนดที่ระดับความเชื่อมั่นมากกว่า 50% จำนวนคำที่ได้จากโปรแกรมกูเกิลคลาวน์วิชันเอพีไอ มีจำนวนมากที่สุดแต่ค่าคำทับซ้อนไม่สูง เนื่องจากคำที่ได้มาเป็นคำที่ได้มาจากการสร้างคำที่มีความหมายใกล้เคียงกับคำที่มีค่าระดับความเชื่อมั่นสูง หากคำที่มีระดับความเชื่อมั่นสูงไม่ถูกต้องตรงกับเมทาดาทาที่ได้มาพร้อมภาพถ่ายจะทำให้คำอื่น ๆ ไม่ถูกต้องไปด้วย แต่ข้อดี

ของการสร้างคำที่มีความหมายใกล้เคียงจำนวนมากจะช่วยเพิ่มประสิทธิภาพคำค้นได้เป็นอย่างดี ทำให้มีค่าคำไม่ซ้อนทับแต่เป็นคำที่ช่วยเพิ่มโอกาสในการสืบค้นคำที่ใกล้เคียงสูง อีกทั้งโปรแกรมยังสามารถดึงข้อความที่อยู่ในรูปภาพ และระบุสถานที่สำคัญต่าง ๆ ได้ดี รวมถึงการเชื่อมโยงไปยังบริการแผนที่กูเกิล (Google map) ให้ได้อีกด้วย

จากผลเมทาดาทาตัวอย่างภาพวิวทิวทัศน์ โปรแกรมกูเกิลคลาวน์วิชันเอพีไอ สามารถระบุคำที่แสดงถึงลักษณะของภูมิประเทศได้อย่างถูกต้อง เช่น “sky” “mountain” “cloud” ได้ในระดับความเชื่อมั่นสูง ในขณะที่โปรแกรมอเมซอน เรคคอคนิชันกำหนดเมทาดาทาระดับความเชื่อมั่นสูงเป็นส่วนประกอบของใบหน้า เช่น “nature” “range” “outdoor” ซึ่งบอกคุณลักษณะแบบกว้างๆ



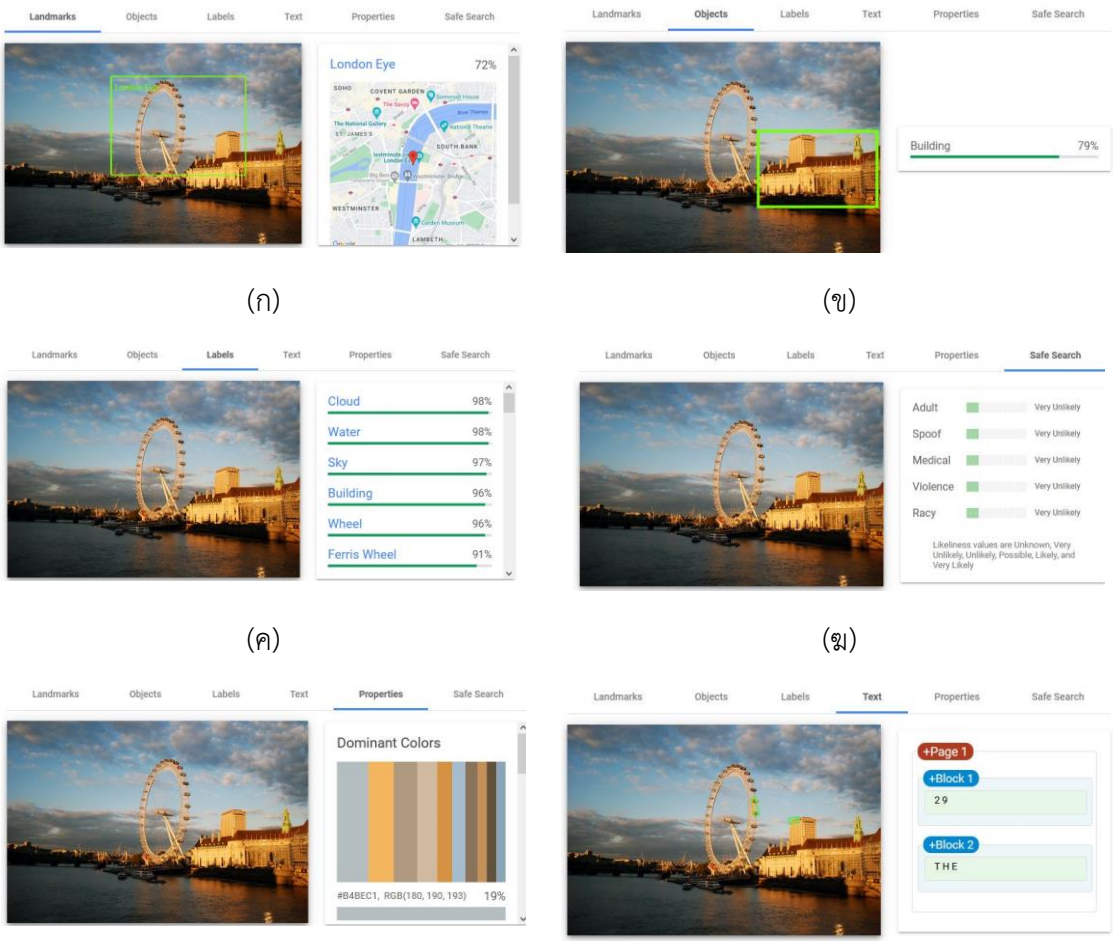
(ก)



(ข)

ภาพที่ 46 เปรียบเทียบเมทาดาทาภาพวิวทิวทัศน์จากโปรแกรม (ก) กูเกิลคลาวน์วิชันเอพีไอ และ (ข) อเมซอน เรคคอคนิชัน

ข้อได้เปรียบของโปรแกรมอเมซอน เรคคอคนิชัน คือความสามารถในการกำหนดข้อมูลที่หลากหลาย เช่น ตำแหน่งสถานที่ (landmark) วัตถุ (object) เมทาดาทา (label) ตัวอักษร (text) คุณสมบัติ (property) และ ความเหมาะสม (safe search) ดังแสดงในภาพที่ 47



(ก) (ข)

(ค) (ฆ)

(ง) (จ)

ภาพที่ 47 ผลการใช้โปรแกรมกูเกิลคลาวน์วิชันเอพีไอกับภาพสถานที่ท่องเที่ยวลอนดอนอาย (The London eye) (ก) ตำแหน่งสถานที่ (ข) วัตถุ (ค) เมทาดาทา (ฆ) ตัวอักษร (ง) คุณสมบัติ และ (จ) ความเหมาะสม

โดยเฉพาะความสามารถในการ แยกตัวอักษรจากภาพ (Text) ดังแสดงในภาพที่ 48 ที่สามารถระบุชื่อป้ายรถเมล์ “Greenwich 177” ในภาพรถเมล์ หรือคำว่า “Mellin’s Food” ในภาพเด็กได้ผลค่อนข้างดี ซึ่งหากเอาคุณสมบัติเหล่านี้มารวมกับเมทาดาทาที่กำหนดจะช่วยเพิ่มประสิทธิภาพของเมทาดาทาได้อย่างมาก

ภาพที่ 48 ผลลัพธ์การใช้โปรแกรมกูเกิลคลาวน์วิชันในการแยกข้อความจากภาพ

ความสามารถกำหนดเมทาดาทารูปภาพโดยอัตโนมัติด้วยเครื่องมือปัญญาประดิษฐ์นี้สามารถช่วยประหยัดเวลา เพิ่มประสิทธิภาพเมทาดาทาได้ โดยเฉพาะผู้ใช้ที่ต้องการจัดระเบียบและค้นหาภาพผ่านคอลเล็กชันรูปภาพจำนวนมาก หากสรุปข้อดีและข้อจำกัดของเครื่องมือแต่ละตัวคือ อเมซอน เรคคอคนิชัน สามารถระบุวัตถุและฉากต่างๆ ภายในภาพได้อย่างแม่นยำ ประสิทธิภาพโดยรวมมีความสม่ำเสมอสูงในชุดทดสอบต่างๆ ซึ่งบ่งชี้ถึงความทนทานในระดับสูง และมีความสามารถในการตรวจจับข้อความและการจดจำใบหน้า คลาริฟายมีความแม่นยำสูงเมื่อต้องระบุวัตถุเฉพาะภายในภาพ และมีความสามารถในการทำความเข้าใจบริบทของรูปภาพและระบุแนวคิดที่เกี่ยวข้อง อิมเมกกามีความโดดเด่นในด้านความสามารถในการแยกและจัดหมวดหมู่ข้อมูลเมตาของรูปภาพได้อย่างแม่นยำ มีคุณสมบัติการครอบตัดรูปภาพและการดึงสี กูเกิลคลาวน์วิชันเอพีไอมีความสามารถในการวิเคราะห์ภาพได้อย่างรวดเร็วและแม่นยำ สามารถตรวจจับวัตถุ การจดจำใบหน้า และการติดฉลากรูปภาพ และสามารถนำไปใช้ร่วมกับบริการอื่น ๆ ของกูเกิลได้ด้วย เช่นการแปลภาษา (Google translate) แผนที่ (Google map) การประมวลผลคำธรรมชาติ (National language) นอกเหนือจากความสามารถในการกำหนดเมทาดาทาแล้วเครื่องมือปัญญาประดิษฐ์แต่ละตัวยังมีความสามารถอื่น ๆ ที่สามารถนำมาใช้ร่วมเพื่อกำหนดเมทาดาทาให้ได้ครบถ้วนสมบูรณ์มากยิ่งขึ้นด้วย ดังนี้

1. อเมซอน เรคคอคนิชัน (Amazon Rekognition) สามารถตรวจจับป้ายชื่อ ตรวจจับใบหน้าและตรวจจับข้อความ ตรวจจับเนื้อหาที่ไม่เหมาะสม จดจำดาราหรือสถานที่ที่นิยมตรวจจับและจดจำข้อความในหลายภาษาได้

2. คลาริฟาย (Clarifai) สามารถตรวจจับป้ายชื่อ ตรวจจับใบหน้า ตรวจจับข้อมูลที่กำหนดเองเช่นสินค้าหรือโลโก้ที่เจาะจง จดจำดาราและสถานที่ที่นิยม และตรวจจับและจดจำข้อความในหลายภาษาได้

3. อิมเมกกา (Imagga) สามารถตรวจจับข้อมูลที่กำหนดเองเช่นสินค้าหรือโลโก้ที่เจาะจง สามารถจดจำดาราและสถานที่ที่นิยม ตรวจจับและจดจำข้อความในหลายภาษาได้

4. กูเกิลคลาวน์วิชันเอพีไอ (Google cloud vision API) สามารถตรวจจับใบหน้า ข้อความตรวจจับเนื้อหาที่ไม่เหมาะสม สามารถจดจำดาราและสถานที่ที่นิยม ตรวจจับและจดจำข้อความในหลายภาษาได้

จากการศึกษาแสดงให้เห็นถึงการใช้โปรแกรมที่เหมาะสมสำหรับภาพถ่ายแต่ละประเภท แต่ในกรณีที่ในภาพถ่ายหนึ่งภาพประกอบด้วยวัตถุหลายประเภท การเลือกใช้เครื่องมือต้องพิจารณาถึงองค์ประกอบหลักของภาพ หรืออาจจะต้องใช้เครื่องมือมากกว่า 1 โปรแกรมทำงานร่วมกันยกตัวอย่างเช่น

- ภาพถ่ายที่ประกอบด้วยคน สัตว์ และ วิวทิวทัศน์ อาจต้องใช้โปรแกรมคลาริฟายที่ทำงานได้ดีในภาพคน สัตว์ ร่วมกับโปรแกรมกูเกิลคลาวน์วิชันเอพีไอที่ทำงานได้ดีในภาพวิวทิวทัศน์
- ภาพถ่ายที่ประกอบด้วยอาหาร และผักผลไม้ อาจต้องใช้โปรแกรมคลาริฟายที่ทำงานได้ดีในภาพอาหาร ร่วมกับโปรแกรมอเมซอน เรคคอคนิชันที่ทำงานได้ดีในภาพผักผลไม้
- ภาพถ่ายที่ประกอบด้วยคน และสถานที่ท่องเที่ยว อาจต้องใช้โปรแกรมคลาริฟายที่ทำงานได้ดีในภาพคน ร่วมกับโปรแกรมกูเกิลคลาวน์วิชันเอพีไอที่ทำงานได้ดีในภาพสถานที่ท่องเที่ยว เป็นต้น

ถึงแม้ว่าจากการศึกษาจะนำเสนอโปรแกรมที่เหมาะสมกับภาพถ่ายแต่ละประเภท การนำไปใช้งานจริงอาจจะต้องใช้ผลจากหลายโปรแกรมประกอบกัน อีกทั้งการพิจารณาหมวดหมู่ภาพถ่ายว่าควรเลือกใช้โปรแกรมใด และการตรวจสอบในลำดับสุดท้ายด้วยมนุษย์ยังเป็นขั้นตอนที่จำเป็น เพื่อให้ได้เมทาดาทาที่ถูกต้องสมบูรณ์มากที่สุด

ข้อเสนอแนะ

จากข้อค้นพบของงานวิจัยมีข้อเสนอแนะที่สามารถนำไปใช้ในการทำงานในอนาคตเพื่อเพิ่มความเชื่อมั่นของการใช้เครื่องมือปัญญาประดิษฐ์ในการกำหนดเมทาดาทาได้ดังนี้

1. การทดลองประเมินประสิทธิภาพของเครื่องมือในโดยกำหนดประเภทของภาพที่มากขึ้น การทดสอบเครื่องมือบนชุดข้อมูลที่ใหญ่ขึ้น และการใช้ภาพประเภทและความละเอียดที่แตกต่างกัน การเปรียบเทียบเครื่องมือในเชิงลึกมากขึ้นโดยใช้วิธีการประเมินแบบอื่น ๆ ร่วมด้วย ตัวอย่างเช่น การประเมินเครื่องมือตามความสามารถในการตรวจจับวัตถุหรือฉากเฉพาะในภาพ หรือการประเมินเมทาดาทาโดยพิจารณาจากความหมายของคำที่มีบริบทใกล้เคียงเพิ่มเติม
2. ศึกษาเครื่องมือปัญญาประดิษฐ์เพิ่มเติมเพื่อให้ได้ข้อสรุปที่ช่วยให้ผู้ใช้งานมีทางเลือกที่หลากหลายมากยิ่งขึ้น
3. เพื่อให้ได้ความเข้าใจที่ครอบคลุมมากขึ้นเกี่ยวกับประสิทธิภาพของเครื่องมือ การประเมินเครื่องมือในแง่ของความสามารถในการสร้างข้อมูลเมทาดาทาในหลายภาษาจะช่วยสนับสนุนความถูกต้องของการวิเคราะห์เครื่องมือได้
4. เพื่อปรับปรุงประสิทธิภาพของเครื่องมือ อาจต้องพัฒนาเทคนิคการประมวลภาพเบื้องต้นผล หรือการปรับแต่งที่เฉพาะเจาะจงสำหรับรูปภาพบางประเภท เช่น รูปภาพโบราณ หรือรูปภาพที่ชำรุด ไม่ชัดเจน
5. ควรทำการวิจัยเพิ่มเติมเกี่ยวกับความคุ้มค่าของเครื่องมือ รวมถึงค่าใช้จ่ายในการใช้เครื่องมือและค่าใช้จ่ายในการบำรุงรักษาและปรับปรุงเครื่องมือเมื่อเวลาผ่านไป
6. พิจารณาความเป็นส่วนตัวและความปลอดภัยของข้อมูล และการผสานรวมกับระบบอื่น ๆ ซึ่งเป็นสิ่งสำคัญในการประเมินในกรณีการใช้งานจริง
7. การดำเนินการวิเคราะห์เชิงลึกเพิ่มเติมของคะแนนความเชื่อมั่นที่ได้รับจากเครื่องมือ
8. การพัฒนาวิธีที่มีประสิทธิภาพมากขึ้นในการเปรียบเทียบประสิทธิภาพของเครื่องมือเพื่อให้การเปรียบเทียบมีความแม่นยำและยุติธรรมมากขึ้น วิธีหนึ่งในการทำเช่นนี้คือการใช้ชุดข้อมูลมาตรฐาน และเมทาดาทาที่กำหนดโดยผู้เชี่ยวชาญทางด้านภาษาโดยเฉพาะ
9. วิเคราะห์ความสามารถในด้านการตรวจจับข้อความในภาพถ่าย โทนสี อารมณ์ ความรู้สึก รวมถึงการระบุบุคคลในภาพ
10. เปรียบเทียบคุณสมบัติในด้านอื่น ๆ ประกอบด้วย เช่น การใช้งานง่าย ความสามารถในการปรับขนาดการให้บริการ การสนับสนุนลูกค้า การรวมเข้ากับเครื่องมืออื่น การจัดทำเอกสารประกอบการใช้งาน และบทวิจารณ์จากลูกค้า

โดยสรุป ยังมีประเด็นที่น่าสนใจสำหรับการวิจัยในอนาคตที่สามารถช่วยปรับปรุงประสิทธิภาพของเครื่องมือปัญญาประดิษฐ์ในการกำหนดเมทาดาทา และเพื่อทำความเข้าใจจุดแข็งและจุดอ่อนของเครื่องมือแต่ละตัวได้ดียิ่งขึ้น อีกทั้งยังเพิ่มความเชื่อมั่นในการใช้เครื่องมือต่าง ๆ ให้ถูกต้องตรงตามวัตถุประสงค์การใช้งานของผู้ใช้ได้มากที่สุด

เอกสารอ้างอิง

ดุษฎีพร ชาติบุตร. (2562). มาตรฐานเมทาดาทาสำหรับพิพิธภัณฑ์: ก้าวใหม่ของพิพิธภัณฑ์ไทยในยุคดิจิทัล. *วารสารสังคมวิทยามานุษยวิทยา, 38*(1), 109-134.

ปราชญ์ สงวนศักดิ์. (2559). การพัฒนาระบบเมทาดาทาเชิงสหสัมพันธ์สู่การบูรณาการด้านนวัตกรรม. *PULINET Journal, 3*(3), 33-42.

วิศปัตย์ ชัยช่วย. (2556). *การพัฒนาเค้าร่างเมทาดาตาสำหรับการจัดการจารึกที่อยู่ในรูปดิจิทัล.* วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต, สาขาวิชาสารสนเทศศึกษา, มหาวิทยาลัยขอนแก่น.

สิทธิชัย บวชไธสง. (2560). ฐานข้อมูลสารสนเทศดิจิทัล วัฒนธรรมและภูมิปัญญาท้องถิ่น จังหวัดศรีสะเกษ. *รังสิตสารสนเทศ, 23*(2), 99-115.

สุวรรณี ห้วยหงส์ทอง และกันยารัตน์ เควียเซ่น. (2562). การกำหนดองค์ประกอบเมทาดาทาสำหรับภาพจิตรกรรมฝาผนัง. *วารสารสารสนเทศศาสตร์, 37*(2), 94-114.

สำนักงานนวัตกรรมแห่งชาติ องค์การมหาชน. (2564). *Aritech.* เข้าถึงได้จาก https://www.nia.or.th/quote-NIAnatomy-ep4

สำนักสถิติแห่งชาติ. (2564). *สำรวจการมีการใช้เทคโนโลยีสารสนเทศและการสื่อสารในครัวเรือน.* เข้าถึงได้จาก http://nso.go.th/sites/2014/DocLib13/ด้านICT/เทคโนโลยีในครัวเรือน/2564/fullreport_ict_q4_64.pdf

Abdullah, S., Sura, M., & Makttof, M. (2020). Modifying jaccard coefficient for texts similarity. *Opcion, 32*, 2899-2921.

Addapa, A., Ramaswamy, P., & Mungara, K. (2020). Object detection/Recognition using machine learning techniques in AWS. *The International Journal of Analytical and Experimental Modal Analysis, XII*(III), 2361-2369.

Amazon Web Services, Inc. (2022). *Amazon rekognition*. Retrieved from https://aws.amazon.com/th/rekognition/

Bakker, R., Rowan, K., Hu, L., Guan, B., Liu, P., Li, Z., He, R., Monge, C. (2020). *AI for archives: Using facial recognition to enhance metadata*. Retrieved from Florida International Libraries: https://digitalcommons.fiu.edu/glworks/93/

Baum, S., & Owe, A. (2022). From AI for people to AI for the world and the universe. *AI & SOCIETY*, 1-2. doi:10.1007/s00146-022-01402-5

Binmakhashen, G., & Mahmoud, S. (2019). Document layout analysis: A comprehensive survey. *ACM Computing Surveys, 52*, 1-36. doi: 10.1145/3355610

Bunn, J. (2021). *Born digital archive cataloguing and description*. Retrieved from https://www.dpconline.org/docs/technology-watch-reports/2468-born-digital-archive-cataloguing-and-description/file

Burke, M., & Zavalina, O. (2020). Descriptive richness of free-text metadata: A comparative analysis of three language archives. *Proceedings of the Association for Information Science and Technology, 57*. doi:10.1002/pra2.429

Büttner, G. (2019). Auto-classification in an international organization: report from a feasibility study. *Comma 2017*(2), 15-26. doi:10.3828/comma.2017.2.2

Cameron, F. (2004). Digital futures II: Museum collections, documentation, and shifting knowledge paradigms. *Collections, 1*, 243-260. doi:10.1177/155019060400100304

Clarifai, Inc. (2022). *Clarifai*. Retrieved from https://www.clarifai.com/

Colavizza, G., Blanke, T., Jeurgens, C., & Noordegraaf, J. (2021). Archives and AI: An overview of current debates and future perspectives. *Journal on Computing and Cultural Heritage, 15*(1), 2-15.

Exif data. (2022). *What is EXIF data.* Retrieved from https://exifdata.com

Gangwar, D. P. G., & Pathania, A. (2018). Authentication of digital image using Exif metadata and decoding properties. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 335-341. doi:10.32628/CSEIT183815

Ghazali, J., Mohammad Niyaz Khan, S., & Zakaria, L. Q. (2020). Image classification using EXIF metadata. *International Journal of Engineering Trends and Technology*, 69-73. doi:10.14445/22315381/CATI3P211

Google cloud platform. (2022) Vision AI. Retrieved from https://cloud.google.com/vision/

Han, H., Giles, C. L., Manavoglu, E., Zha, H., Zhang, Z., & Fox, E. (2003). *Automatic document metadata extraction using support vector machines*. Paper presented at the 2003 Joint Conference on Digital Libraries Proceeding. (27-31 May 2003) Houston, TX, USA: IEEE.

Hancock, J. (2004). Jaccard Distance (Jaccard Index, Jaccard Similarity Coefficient). In *Dictionary of bioinformatics and computational biology*. Colchester: John Wiley & Sons

Harran, M., Farrelly, W., & Curran, K. (2017). A method for verifying integrity & authenticating digital media. *Applied Computing and Informatics, 14*. doi:10.1016/j.aci.2017.05.006

Hong, S.-Y., & Lee, S.-J. (2015). An intelligent web digital image metadata service platform for social curation commerce environment. *Modelling and Simulation in Engineering, 2015*, 651428. doi:10.1155/2015/651428

Imagga Technologies. (2022) *Imagga.* Retrieved from https://imagga.com/

International Press Telecommunications Council. (2022). *Photo metadata*. Retrieved from https://www.iptc.org/standards/photo-metadata/

Jaffe, R. (2020). Rethinking metadata's value and how it is evaluated. *Technical Services Quarterly, 37*, 432-443. doi:10.1080/07317131.2020.1810443

Jaillant, L., & Caputo, A. (2022). Unlocking digital archives: cross-disciplinary perspectives on AI and born-digital data. *AI & SOCIETY*. doi:10.1007/s00146-021-01367-x

Kim, H., & Lee, H.-K. (2022). Emotions and colors in a design archiving system: Applying AI technology for museums. *Applied Sciences, 12*, 2467. doi:10.3390/app12052467

Kuzma, M., & Moscicka, A. (2020). Metadata evaluation criteria in respect to archival maps description: A systematic literature review. *The Electronic Library, ahead-of-print*. doi:10.1108/EL-07-2019-0161

Kyriakou, K., Barlas, P., Kleanthous, S., & Otterbacher, J. Fairness in Proprietary Image Tagging Algorithms: A Cross-Platform Audit on People Images. Proceedings of the Thirteenth International AAAI Conference on Web and Social Media; 2019 June 11-14; Münich Germany; 2019. pp. 313-322.

Lee, B. (2019). Machine learning, template matching, and the International Tracing Service digital archive: Automating the retrieval of death certificate reference cards from 40 million document scans. *Digit. Scholarsh. Humanit, 34*, 513-535.

Li, L., & Fleischmann, K. (2020). Libraries and archives of tomorrow: How future information professionals perceive AI. *Proceedings of the Association for Information Science and Technology, 57*. doi:10.1002/pra2.343

Mishra, A. (2019). Amazon Rekognition. In *In book: Machine learning in the AWS Cloud*. New York: John Wiley & Sons.

Muehlberger, G., Seaward, L., Terras, M., Ares Oliveira, S., Bosch, V., Bryan, M., Colutto, S., Déjean, H., Diem, M., Fiel, S., Gatos, B., Greinoecker, A., Grüning, T., Hackl, G., Haukkovaara, V., Heyer, G., Hirvonen, L., Hodel, T., Jokinen, M., Kahle, P., Kallio, M., Kaplan, F., Kleber, F., Labahn, R., Lang, E.M., Laube, S., Leifert, G., Louloudis, G., McNicholl, R., Meunier, J.-L., Michael, J., Mühlbauer, E., Philipp, N., Pratikakis, I., Puigcerver Pérez, J., Putz, H., Retsinas, G., Romero, V., Sablatnig, R., Sánchez, J. A., Schofield, P., Sfikas, G., Sieber, C., Stamatopoulos, N., Strauß, T., Terbul, T., Toselli, A. H., Ulreich, B., Villegas, M., Vidal, E., Walcher, J., Weidemann, M., Wurster, H., & Zagoris, K. (2019). Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study. *Journal of Documentation, ahead-of-print*, 75(5), 954-976. doi:10.1108/JD-07-2018-0114

Mustafa, G., Usman, M., Afzal, M. T., Shahid, A., & Koubaa, A. (2021). A comprehensive evaluation of metadata-based features to classify research paper's topics. *IEEE Access, 9*, 133500-133509. doi:10.1109/ACCESS.2021.3115148

Nichols, D., Chan, C.-H., Bainbridge, D., McKay, D., & Twidale, M. (2022). A tool for metadata analysis. Retrieved from The University of Waikato, Hamilton, New Zealand.

Ochoa, X., & Duval, E. (2009). Automatic evaluation of metadata quality in digital libraries. *International Journal on Digital Libraries, 10*, 67-91. doi:10.1007/s00799-009-0054-4

Rahaman, W. (2017). Digital libraries, digital repositories, digital copyright: Overview, challenges and solutions in technology Era. In S. Pal Kumar & B. Ram (Eds.), *Applications of modern tools and technology in library services* (pp. 157-162). New Delhi: Studera Press.

Rees, A. (2020). Image Recognition as a Tool in Cataloguing Born-Digital Photography. In *Connect to collect: approaches to collecting social digital photography in museums and archives* (pp. 219-231). Stockholm: The Nordic Museum.

Roszkiewicz, R. (2005). Upstream metadata and the impact of the IPTC schema for XMP and Adobe Bridge. *Journal of Digital Asset Management, 1*, 298-302. doi:10.1057/palgrave.dam.3640047

Singthongchai, J., & Niwattanakul, S. (2013). A method for measuring keywords similarity by applying Jaccards, N-Gram and vector space. *Information Theory, 1*(4), 159-164. doi:10.12720/lnit.1.4.159-164

Steele, T., & Sump-crethar, N. (2016). Metadata for electronic theses and dissertations: A survey of institutional repositories. *Journal of Library Metadata, 16*, 53-68. doi:10.1080/19386389.2016.1161462

Steen, N. (2014). Metadata management in Europeana Photography. *SCIRES-IT: SCIentific RESearch and Information Technology, 4*. doi:10.2423/i22394303v4n2p127

Sushma, C., & Kumari, P. L. (2021). Face Recognition using CNN and Tracking Locations using Image Metadata. *Journal of Physics: Conference Series, 2040*, 012009. doi:10.1088/1742-6596/2040/1/012009

Tao, C., Gao, J., & Wang, T. (2019). Testing and Quality Validation for AI Software–Perspectives, Issues, and Practices. *IEEE Access, 7*, 120164-120175. doi:10.1109/ACCESS.2019.2937107

Tesic, J. (2005). Metadata practices for consumer photos. *IEEE MultiMedia, 12*(3), 86-92. doi:10.1109/MMUL.2005.50

Witten, I. H., Bainbridge, D., & Nichols, D. M. (2010). Chapter 6 - Metadata: Elements of organization. In I. H. Witten, D. Bainbridge, & D. M. Nichols (Eds.), *How to build a digital library* (2nd ed.) (pp. 285-341). Boston: Morgan Kaufmann.

Yang, T., Zhao, X., Wang, X., & Lv, H. (2020). Evaluating facial recognition web services with adversarial and synthetic samples. *Neurocomputing, 406*. doi:10.1016/j.neucom.2019.11.117

Zavalina, O. (2013). Complementarity in subject metadata in large-scale digital libraries: A comparative analysis. *Cataloging & Classification Quarterly, 52*, 77-89. doi:10.1080/01639374.2013.848316

Zavalina, O., Kizhakkethil, P., & Shakeri, S. (2015). Metadata change in traditional library collections and digital repositories: Exploratory comparative analysis. *Proceedings of the Association for Information Science and Technology, 52*(1), 1-5. doi:10.1002/pra2.2015.1450520100146