Received: 28 ก.พ. 2569 **Revised:** 9 เม.ย. 2569 **Accepted:** 17 เม.ย. 2569

การพัฒนาระบบเพื่อการจัดการและสังเคราะห์สื่อวิดีโออัตโนมัติจากเนื้อหาดิจิทัล โดยประยุกต์ใช้เทคโนโลยีเวิร์กโฟลว์

Development of Management and Synthesis for Automated Video from Digital Content System with Application of Workflow Technology

ธัญญลักษ์ สุขเกษม[1], รุ่งฤทธิ์ อนุตรวิรามกุล[2], สมนึก สินธุปวน[1],
กิตติกร หาญตระกูล[1] และ ภานุวัฒน์ เมฆะ[1]*
[1]สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้
[2]ศูนย์ปฏิบัติการหอดูดาวงานวิศวกรรมเมคาทรอนิกส์
สถาบันวิจัยดาราศาสตร์แห่งชาติ
(องค์การมหาชน)

Tanyaluk Sukkasem[1], Rungrit Anutarawiramkul[2], Somnuek Sinthupuan[1],
Kittikorn Hantrakul[1] and Panuwat Mekha[1]*
[1]Department of Computer Science, Faculty of Science, Maejo University
[2]Mechatronics, National Astronomical Research Institute of Thailand
(Public Organization)
*Corresponding author: panutwat_m@mju.ac.th

Abstract

Effective science communication, particularly in astronomy, remains a significant challenge for research institutions due to the resource-intensive nature of video production. Standard dissemination cycles typically demand substantial time and personnel. To address this, an automated system was developed to transform digital content into monthly astronomical summary videos, allowing users to define content descriptions and visual assets for automated processing.

The system integrates core functionalities such as text-to-animation conversion, image composition, and automated publishing workflows. Its technical framework leverages the N8N platform for workflow automation, web servers for content management, and AI-driven image processing. These components work together to streamline the transition from raw digital materials to finalized multimedia outputs.

The researchers implemented this automated workflow to eliminate redundant tasks and enhance production efficiency for astronomical media. Performance benchmarking revealed distinct advantages among AI models: Runway (Model A) excelled in processing speed, averaging 2 minutes at a 1200×780 resolution, while Framepack (Model C) offered moderate resolution at 770×520. Additionally, the F5-TTS-Thai speech synthesis and image models operate as local-based, cost-free solutions. These findings confirm that selecting the optimal model configuration effectively meets institutional requirements while optimizing operational resources.

Keyword: *Workflow automation; N8N; Artificial Intelligence; Astronomy; Outreach video content*

บทคัดย่อ

ปัจจุบันการสื่อสารข้อมูลด้านวิทยาศาสตร์ โดยเฉพาะดาราศาสตร์ในรูปแบบที่เข้าถึงง่ายและน่าสนใจ ยังคงเป็นความท้าทายสำคัญของหน่วยงานด้านการวิจัยดาราศาสตร์ เนื่องจากการผลิตสื่อวิดีโอจำเป็นต้องใช้เวลาและทรัพยากรบุคลากรจำนวนมากในแต่ละรอบการเผยแพร่ข้อมูล ซึ่งระบบที่พัฒนาขึ้นเป็นระบบสร้างและจัดการสื่อวิดีโออัตโนมัติจากเนื้อหาดิจิทัล โดยผู้ใช้สามารถกำหนดคำอธิบายเนื้อหาและจัดเตรียมภาพประกอบ เพื่อให้ระบบสร้างวิดีโอสรุปข้อมูลรายเดือนเกี่ยวกับเหตุการณ์ทางดาราศาสตร์ ระบบประกอบด้วยฟังก์ชันสำคัญ เช่น การแปลงข้อความเป็นภาพเคลื่อนไหว การจัดองค์ประกอบภาพ และการจัดการเวิร์กโฟลว์แบบอัตโนมัติในการเผยแพร่ เทคโนโลยีหลักที่ใช้ ได้แก่ เวิร์กโฟลว์อัตโนมัติผ่านแพลตฟอร์มเอ็นเอทเอ็น การจัดการเนื้อหาผ่านเว็บเซิร์ฟเวอร์ และการประมวลผลภาพด้วยเครื่องมือปัญญาประดิษฐ์

ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะนำเทคโนโลยีเวิร์กโฟลว์อัตโนมัติมาช่วยลดขั้นตอนการทำงานซ้ำซ้อน และเพิ่มประสิทธิภาพการผลิตสื่อวิดีโอที่เกี่ยวข้องกับปรากฏการณ์ทางดาราศาสตร์ โดยผลการทดสอบเปรียบเทียบสมรรถนะของโมเดลปัญญาประดิษฐ์พบว่า การประยุกต์ใช้เทคโนโลยีเวิร์กโฟลว์ร่วมกับโมเดลแต่ละประเภทมีข้อเด่นที่ต่างกัน โดยโมเดล Runway (Model A) มีประสิทธิภาพโดดเด่นด้านความเร็วในการประมวลผลเฉลี่ยประมาณ 2 นาที และให้ความละเอียดสูงสุดที่ 1200×780 พิกเซล ในขณะที่ Framepack (Model C) ให้ความละเอียดในระดับปานกลางที่ 770×520 พิกเซล ส่วน F5-TTS-Thai ในด้านการสังเคราะห์เสียง ทั้งสองโมเดลเป็นระบบประมวลผลภายใน (Local-based) และไม่มีค่าบริการจึงใช้ระยะเวลาในการประมวลผล ซึ่งผลการทดลองทั้งหมดนี้ยืนยันว่าการคัดเลือกโมเดลที่เหมาะสมสามารถผลิตสื่อวิดีโอได้ตรงตามความต้องการของผู้วิจัย

คำสำคัญ : *เวิร์กโฟลว์อัตโนมัติ; เอ็นเอทเอ็น; ปัญญาประดิษฐ์; ดาราศาสตร์; ผลิตสื่อวิดีโอ*

1. บทนำ

ปัจจุบันการเติบโตของเทคโนโลยีดิจิทัลและการเชื่อมต่อทางอินเทอร์เน็ตได้เปลี่ยนแปลงรูปแบบการผลิตและเผยแพร่เนื้อหาความรู้อย่างมีนัยสำคัญ องค์กรทั้งภาครัฐและเอกชนต่างเผชิญกับความท้าทายในการจัดการเนื้อหาดิจิทัลปริมาณมหาศาลที่ต้องการความถูกต้อง รวดเร็ว และต่อเนื่อง จึงทำให้แนวคิดการบริหารจัดการเนื้อหาดิจิทัล (Digital content management) กลายเป็นกลยุทธ์สำคัญขององค์กรยุคใหม่ โดยแนวคิดดังกล่าวให้ความสำคัญกับการจัดการวงจรชีวิตของเนื้อหาตั้งแต่การสร้าง การประมวลผล ไปจนถึงการเผยแพร่อย่างเป็นระบบ ผ่านเทคโนโลยีเวิร์กโฟลว์อัตโนมัติ (workflow automation) ที่ช่วยลดขั้นตอนซ้ำซ้อนและเพิ่มความสามารถในการตอบสนองได้อย่างทันท่วงที โดยเฉพาะอย่างยิ่งเมื่อเนื้อหานั้นมีความซับซ้อนและอ่อนไหวต่อเวลา เช่น ข้อมูลทางดาราศาสตร์และปรากฏการณ์บนท้องฟ้า ซึ่งหากกระบวนการผลิตสื่อมีความล่าช้าอาจทำให้ข้อมูลสูญเสียความน่าสนใจไปตามกาลเวลา

ปัจจัยที่ส่งผลต่อคุณภาพของสื่อวิดีโอประกอบด้วย 4 ส่วนหลัก ได้แก่ 1) ความถูกต้องของภาพ (Visual Fidelity) 2) ความไหลลื่นของภาพเคลื่อนไหว (Temporal Consistency) (Adobe , (n.d.)) 3) คุณภาพของเสียงบรรยาย (Acoustic Quality) ที่ต้องมีความถูกต้องของคำศัพท์เฉพาะทาง และ 4) ประสิทธิภาพในการประมวลผลและต้นทุนทรัพยากร (Computational Efficiency & Cost) เนื่องจากแต่ละระบบมีรูปแบบการประมวลผลที่ต่างกัน ทั้งในลักษณะคลาวด์ (cloud-based) (Hewlett Packard Enterprise, (2025)) และการประมวลผลภายในเครื่อง

การพัฒนาเทคโนโลยี Generative AI ในปัจจุบันได้มุ่งเน้นการประมวลผลเพื่อสร้างผลลัพธ์ที่ใกล้เคียงกับมนุษย์ โดยเทคนิคที่ได้รับความสนใจอย่างมากคือ การสร้างวิดีโอจากรูปภาพต้นฉบับ (image-to-video) (Xie, et al. (2025)) ผ่านโมเดล Runway, Wan2.1 และ Framepack ซึ่งอาศัยการประมวลผลร่วมกับชุดคำสั่ง (Prompt) เพื่อกำหนดทิศทางการเคลื่อนไหว (Yan, et al. (2023)) และ การสร้างเสียงบรรยายภาษาไทยจากข้อความ (text-to-speech) (Kaur, N., & Singh, P. (2023)) ผ่านโมเดล Botnoi และ F5-TTS-Thai ซึ่งแต่ละอัลกอริทึมมีความแตกต่างกันทั้งในด้านความละเอียดวิดีโอ อัตราการบีบอัดข้อมูล รวมถึงความเป็นธรรมชาติและความถูกต้องในการออกเสียงคำศัพท์เทคนิคทางดาราศาสตร์

ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะเปรียบเทียบประสิทธิภาพของอัลกอริทึมดังกล่าว โดยกำหนดเกณฑ์มาตรฐานในการทดสอบคือ การใช้รูปภาพต้นฉบับเดียวกันและชุดคำสั่งเดียวกันสำหรับวิดีโอจำนวน 10 ชุดทดสอบ และการใช้ชุดประโยคทดสอบภาษาไทยที่มีคำศัพท์เทคนิคทางดาราศาสตร์จำนวน 10 ประโยคสำหรับเสียงบรรยาย เพื่อวิเคราะห์หาโมเดลที่มีความเหมาะสมที่สุด ทั้งในด้านคุณภาพของผลลัพธ์ และความคุ้มค่าในการบริหารจัดการทรัพยากรของสถาบันวิจัยดาราศาสตร์แห่งชาติ เพื่อให้การเผยแพร่ข้อมูลทางวิทยาศาสตร์เป็นไปได้อย่างรวดเร็ว แม่นยำ และมีประสิทธิภาพสูงสุด

2. วัตถุประสงค์

1.เพื่อเปรียบเทียบคุณภาพของวิดีโอที่สร้างจากรูปภาพต้นฉบับ ผ่านโมเดล Runway, Wan2.1 และ Framepack ในด้านความถูกต้องของภาพ และความไหลลื่นของภาพเคลื่อนไหว

2. เพื่อเปรียบเทียบคุณภาพของเสียงบรรยายภาษาไทยจากข้อความ ผ่านโมเดล Botnoi และ F5-TTS-Thai ในด้านความเป็นธรรมชาติและความถูกต้องในการออกเสียงคำศัพท์เทคนิคทางดาราศาสตร์

3. เพื่อประเมินประสิทธิภาพด้านการประมวลผลและความคุ้มค่าของทรัพยากรของแต่ละโมเดลทั้งในรูปแบบคลาวด์ และการประมวลผลภายในเครื่อง

3. ข้อมูลและวิธีการดำเนินงานวิจัย

การวิจัยนี้มุ่งเน้นเปรียบเทียบประสิทธิภาพของโมเดลปัญญาประดิษฐ์ในการสร้างวิดีโอและสังเคราะห์เสียงภาษาไทย โดยวิเคราะห์ขีดความสามารถเชิงเทคนิคและคุณภาพผลลัพธ์ผ่านระบบเว็บแอปพลิเคชัน

3.1 การศึกษาและวิเคราะห์ระบบ

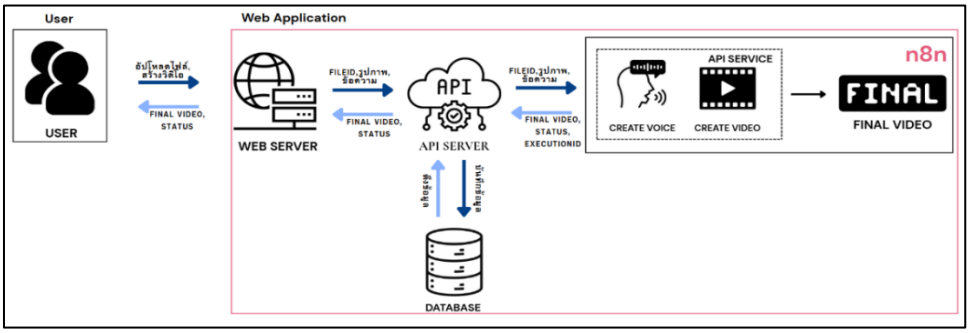
ศึกษาเทคโนโลยีเอ็นเอทเอ็นและการพัฒนาเซิร์ฟเวอร์ เพื่อวางแผนวิเคราะห์และออกแบบระบบให้ตอบโจทย์ผู้ใช้งาน พร้อมจัดทำ Flowchart แสดงลำดับขั้นตอนการทำงานของฟังก์ชันต่างๆ อย่างเป็นระบบ

3.2 วิธีดำเนินการวิจัย

ทดสอบเปรียบเทียบประสิทธิภาพระหว่าง Runway, Wan2.1 และ Framepack ในด้านการสร้างวิดีโอ และระหว่าง Botnoi กับ F5-TTS-Thai ในด้านการสังเคราะห์เสียง โดยประเมินจากคุณลักษณะทางเทคนิคของไฟล์ควบคู่กับความถูกต้องแม่นยำของเนื้อหา

3.3 สถาปัตยกรรมของระบบ

เว็บแอปพลิเคชันทำหน้าที่เป็นส่วนติดต่อและตัวกลางสื่อสารข้อมูลระหว่างผู้ใช้กับส่วนประมวลผลดังภาพ



ภาพที่ 1 กระบวนการทำงานของระบบ

จากภาพที่ 1 แสดงถึงกระบวนการทำงานของระบบ มีการแบ่งลำดับการทำงานรับ-ส่งข้อมูลออกเป็น 2 ส่วนหลักด้วยกัน ดังต่อไปนี้

1. ผู้ใช้งาน (User) มีหน้าที่ป้อนข้อมูลตั้งต้นและสั่งเริ่มกระบวนการผ่านเว็บแอปพลิเคชันเพื่อส่งข้อมูลเข้าสู่ระบบสร้างวิดีโอ

2. เว็บแอปพลิเคชัน (Web Application) ทำหน้าที่เป็นตัวกลางสื่อสารข้อมูลและนำเสนอผลลัพธ์แก่ผู้ใช้ โดยมีองค์ประกอบหลักดังนี้

2.1 ส่วนติดต่อผู้ใช้งาน (Frontend) พัฒนาด้วย Next.js และ Tailwind CSS เพื่อสร้าง User Interface ที่ตอบสนองและเชื่อมต่อกับส่วนหลังบ้านผ่าน API

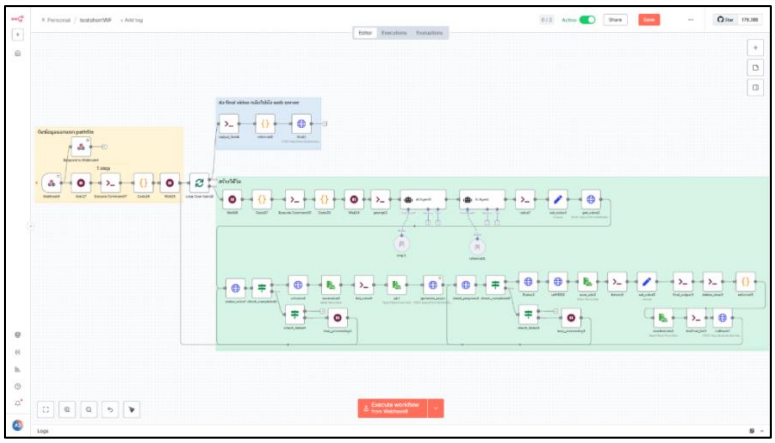
2.2 ส่วนจัดการข้อมูล (Backend) พัฒนาด้วย Node.js (TypeScript) ทำหน้าที่จัดการตรรกะทางธุรกิจ การตรวจสอบสิทธิ์ และสื่อสารกับระบบเวิร์กโฟลว์

2.3 ระบบจัดการฐานข้อมูล ใช้ MongoDB จัดเก็บข้อมูลรูปแบบ BSON เพื่อความรวดเร็วและมีประสิทธิภาพ (ผุสดี บุญรอด และประกายมาศ ศรีสุขทักษิณ, (2015)) ประกอบด้วย 5 ตารางหลัก

2.4 ระบบเวิร์กโฟลว์อัตโนมัติ ใช้เอ็นเอทเอ็นบน Docker (Venkiteela, P., (n.d.)) เพื่อประสานงานระหว่าง AI Models และ FFmpeg ในการสร้างสื่อวิดีโอ

3.4 การออกแบบเวิร์กโฟลว์

เวิร์กโฟลว์บนเอ็นเอทเอ็นถูกออกแบบให้ทำงานอัตโนมัติแบบครบวงจร เริ่มจากการรับข้อมูลผ่าน Webhook เพื่อเข้าสู่กระบวนการประมวลผลไฟล์และส่งต่อไปยัง AI Models สำหรับสร้างวิดีโอ จากนั้นจึงดำเนินการรวมสื่อด้วย FFmpeg และบันทึกผลลัพธ์พร้อมแจ้งเตือนกลับผ่าน Callback API



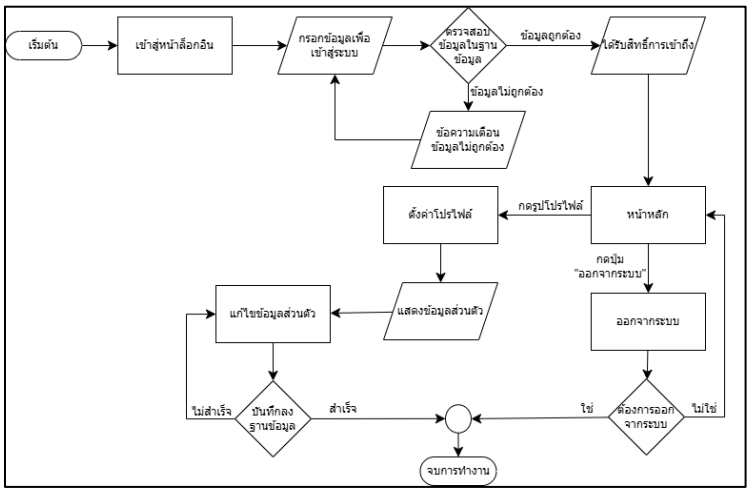
ภาพที่ 2 แสดงโครงสร้างเวิร์กโฟลว์บนเอ็นเอทเอ็นที่ได้ออกแบบไว้

จากภาพที่ 2 เวิร์กโฟลว์อัตโนมัติบนเอ็นเอทเอ็นแบ่งการทำงานเป็น 3 ส่วนหลัก ได้แก่ ส่วนที่ 1 การรับและจำแนกข้อมูลนำเข้า (trigger) เพื่อเตรียมการผลิต ส่วนที่ 2 การสร้างวิดีโอย่อยแบบวน

ซ้ำ (loop) พร้อมส่งสถานะกลับแบบเรียลไทม์ และส่วนที่ 3 การรวบรวมวิดีโอย่อยเป็นไฟล์ฉบับสมบูรณ์ (final video) เพื่อส่งมอบผลลัพธ์แก่ผู้ใช้งานผ่านเว็บแอปพลิเคชัน

3.5 การพัฒนาระบบจัดการผู้ใช้งาน

ระบบจัดการผู้ใช้งานมีฟังก์ชันหลัก ได้แก่ การเข้าสู่ระบบ การจัดการข้อมูลส่วนตัว และการจัดการโทเคนสำหรับควบคุมการสร้างวิดีโอ เจ้าหน้าที่สามารถเพิ่มหรือปรับโทเคนของผู้ใช้ได้ ระบบมีการเข้ารหัสรหัสผ่าน ตรวจสอบสิทธิ์การเข้าถึง และบันทึกประวัติการใช้งานเพื่อตรวจสอบย้อนหลัง

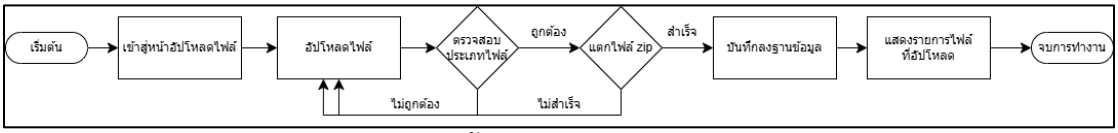


ภาพที่ 3 ขั้นตอนการทำงานของระบบจัดการผู้ใช้งาน

จากภาพที่ 3 แสดงขั้นตอนการทำงานของระบบ เริ่มจากการเข้าสู่ระบบและตรวจสอบข้อมูล โดยแบ่งสิทธิ์เป็นผู้ใช้ทั่วไปและเจ้าหน้าที่พร้อมบันทึกทุกการดำเนินการไว้ตรวจสอบ

ระบบรองรับการอัปโหลดไฟล์ .zip ที่มีรูปภาพ คำสั่งสร้างวิดีโอ และข้อความเสียงบรรยาย โดยตรวจสอบ แตกไฟล์ และบันทึกลงฐานข้อมูลโดยอัตโนมัติ ผู้ใช้สามารถสร้างวิดีโอ ติดตามสถานะ และดาวน์โหลดผลลัพธ์ได้

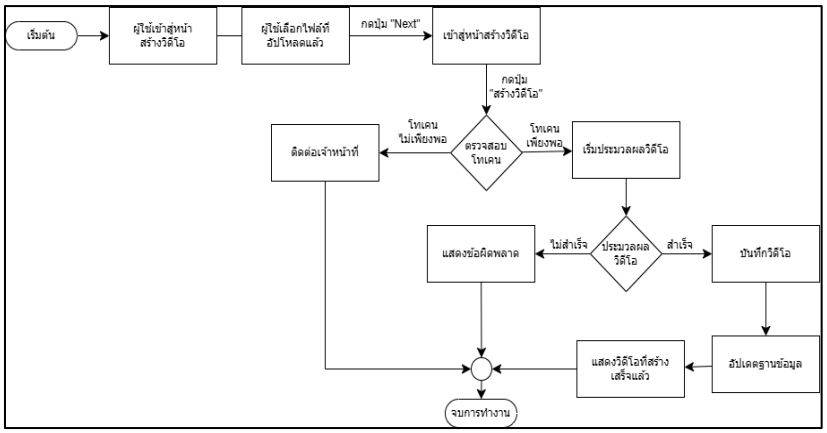
1. ขั้นตอนการอัปโหลดและจัดการไฟล์



ภาพที่ 4 แผนภาพขั้นตอนการอัปโหลดและจัดการไฟล์

จากภาพที่ 4 ระบบตรวจสอบรูปแบบและขนาดไฟล์ หากไม่ผ่านจะแจ้งเตือนให้เลือกใหม่ เมื่อผ่านแล้วจะอัปโหลด แตกไฟล์ และบันทึกลงฐานข้อมูล จากนั้นผู้ใช้เลือกไฟล์เพื่อสร้างวิดีโอได้

2. ขั้นตอนการสร้างวิดีโอ



ภาพที่ 5 แผนภาพขั้นตอนการสร้างวิดีโอ

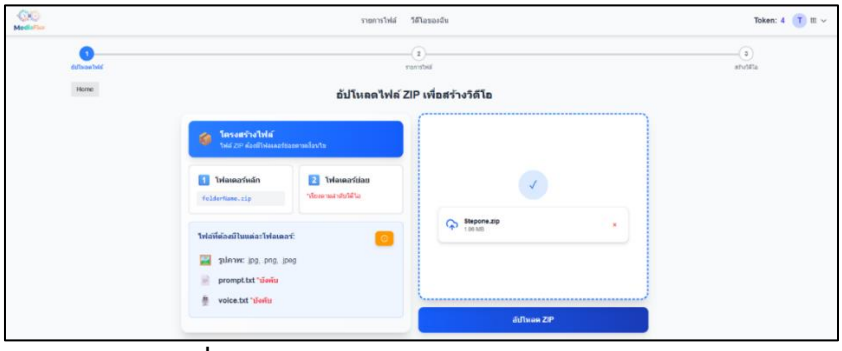
จากภาพที่ 5 ระบบตรวจสอบไฟล์และโทเคนก่อน หากผ่านจะตรวจสอบวิดีโอเดิมแล้วสร้างหรือแสดงผลตามลำดับ ผู้ใช้ดาวน์โหลดได้ทันทีหรือติดตามสถานะผ่านหน้าประวัติ

3.6 การออกแบบส่วนติดต่อผู้ใช้งานและกระบวนการทำงานของระบบ

ผู้วิจัยออกแบบส่วนติดต่อผู้ใช้งานโดยเน้นความง่ายและครบถ้วนของฟังก์ชันหลักดังนี้

1. ส่วนการนำเข้าข้อมูลและจัดการโครงสร้างไฟล์

ระบบรองรับการนำเข้าข้อมูลในรูปแบบไฟล์ .zip โดยผู้ใช้ต้องจัดเตรียมโครงสร้างข้อมูลแบบแยกโฟลเดอร์ย่อย

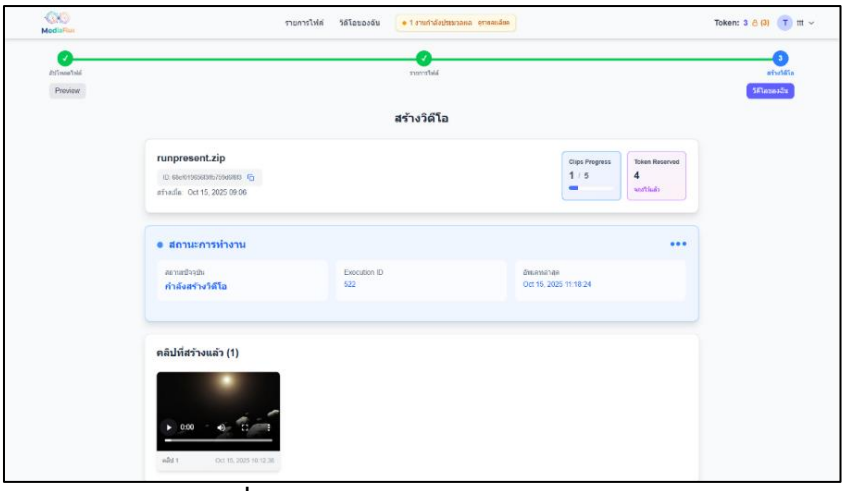


ภาพที่ 6 การนำเข้าข้อมูลและจัดการโครงสร้างไฟล์

จากภาพที่ 6 แสดงโครงสร้างไฟล์ .zip ที่แบ่งเป็นโฟลเดอร์ย่อย โดยแต่ละโฟลเดอร์แทน 1 ฉาก ประกอบด้วยไฟล์ภาพ prompt.txt และ voice.txt ตามลำดับ

2. ส่วนติดตามและแสดงผลการประมวลผล

ส่วนนี้ออกแบบมาเพื่อให้ผู้ใช้สามารถติดตามสถานะของระบบได้แบบเรียลไทม์

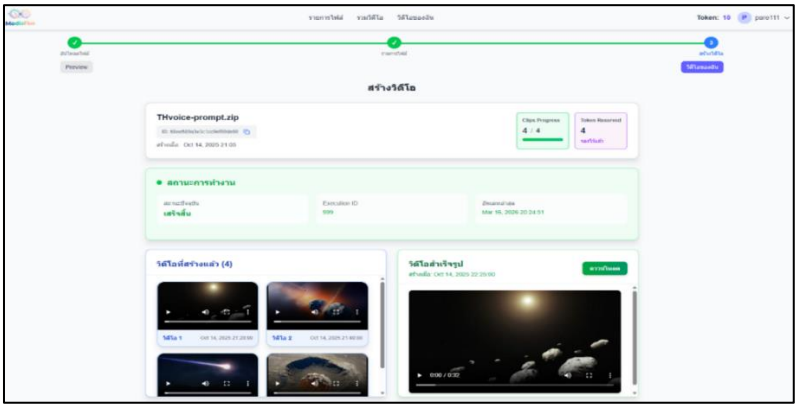


ภาพที่ 8 การติดตามความคืบหน้าแต่ละฉาก

จากภาพที่ 8 แสดงหมายเลขและขั้นตอนการประมวลผลปัจจุบัน เมื่อแต่ละฉากเสร็จสิ้น ระบบจะแสดงคลิปวิดีโอแยกตามฉากเพื่อให้ตรวจสอบก่อนรวมไฟล์

3. ส่วนแสดงผลลัพธ์สุดท้ายและการจัดการวิดีโอ

ระบบรวบรวมคลิปทุกฉากพร้อมเสียงบรรยายเป็นวิดีโอฉบับสมบูรณ์

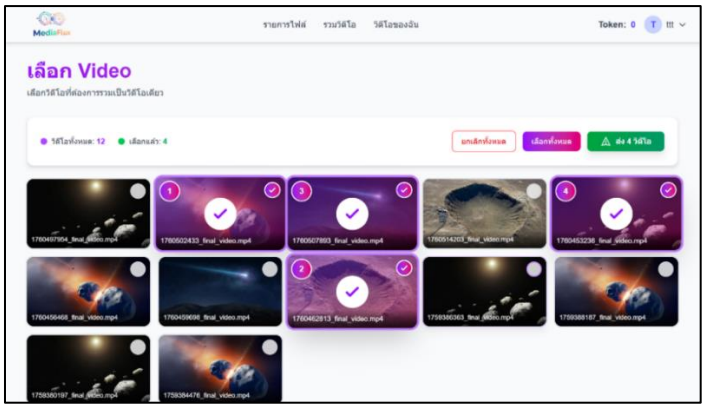


ภาพที่ 9 รวบรวมคลิปวิดีโอจากทุกฉากเป็นวิดีโอฉบับสมบูรณ์

จากภาพที่ 9 แสดงคลิปวิดีโอแต่ละฉากและวิดีโอฉบับสมบูรณ์ที่รวมทุกฉากแล้ว ผู้ใช้งานสามารถตรวจสอบ และดาวน์โหลดได้โดยตรงจากหน้าจอนี้

4. ระบบจัดการคลังทรัพยากรและรวมสื่อ

ระบบรองรับการนำทรัพยากรดิจิทัลกลับมาใช้ซ้ำเพื่อเพิ่มความยืดหยุ่นในการผลิตสื่อ



ภาพที่ 10 การจัดการคลังวิดีโอและการรวมสื่อใหม่

จากภาพที่ 10 ผู้ใช้สามารถเลือกวิดีโอจากคลังมาจัดลำดับและรวมเป็นวิดีโอชุดใหม่ได้ ช่วยลดความซ้ำซ้อนและเพิ่มความรวดเร็วในการผลิตสื่อ

4. ผลการทดลอง

ผู้วิจัยทดสอบเปรียบเทียบระบบ AI Models สร้างวิดีโอ 3 รูปแบบ ได้แก่ Runway (Model A), Wan2.1 (Model B) และ Framepack (Model C) โดยใช้ชุดข้อมูล 10 ชุด ภายใต้รูปภาพและคำสั่งเดียวกัน เพื่อวิเคราะห์ทั้งด้านเทคนิคและคุณภาพของผลลัพธ์ ดังนี้

ตารางที่ 1 การเปรียบเทียบคุณสมบัติทางเทคนิคของวิดีโอ

ตัวอย่างที่	ระบบ AI	ความยาว (วินาที)	ความละเอียดวิดีโอ (W×H)	Total Bitrate (kbps)	Frame Rate (fps)	ขนาดไฟล์ (MB)
S01	Runway	5	960×960	2,028	24	1.9
	Wan2.1	5	944×960	1,404	30	0.9
	Framepack	8	608×640	1,156	30	1.1
S02	Runway	5	1280×720	1,703	24	1.8
	Wan2.1	5	1280×704	3,017	30	0.9
	Framepack	8	832×480	1,300	30	1.0
S03	Runway	5	1280×720	1,570	24	1.9
	Wan2.1	5	1280×704	7,371	30	0.8
	Framepack	8	832×480	1,711	30	1.1
S04	Runway	5	1280×720	3,325	24	1.8
	Wan2.1	5	1280×720	1,652	30	0.9
	Framepack	8	832×480	5,501	30	1.0
S05	Runway	5	1280×720	1,274	24	2.0
	Wan2.1	5	1280×720	824	30	0.8
	Framepack	8	832×480	652	30	1.1
S06	Runway	5	960×960	2,756	24	1.9
	Wan2.1	5	960×960	2,746	30	0.9
	Framepack	8	608×640	4,491	30	1.1
S07	Runway	5	1280×720	1,535	24	1.8
	Wan2.1	5	1200×752	634	30	0.9
	Framepack	8	768×512	2,660	30	1.0
S08	Runway	5	1104×832	5,753	24	1.9
	Wan2.1	5	1152×784	3,576	30	0.9
	Framepack	8	768×512	2,169	30	1.1

ตารางที่ 1 การเปรียบเทียบคุณสมบัติทางเทคนิคของวิดีโอ (ต่อ)

ตัวอย่างที่	ระบบ AI	ความยาว (วินาที)	ความละเอียดวิดีโอ (WxH)	Total Bitrate (kbps)	Frame Rate (fps)	ขนาดไฟล์ (MB)
S09	Runway	5	1280×720	4,547	24	1.9
	Wan2.1	5	1280×720	2,856	30	0.9
	Framepack	8	832×480	3,578	30	1.1
S10	Runway	5	1280×720	3,265	24	1.8
	Wan2.1	5	1280×720	3,572	30	0.9
	Framepack	8	832×480	4,565	30	1.0

ตารางที่ 2 ค่าเฉลี่ยของคุณสมบัติทางเทคนิค

คุณสมบัติ	ระบบ AI			ค่าที่ดีที่สุด
	Runway	Wan2.1	Framepack	
ความยาวเฉลี่ย (วินาที)	5.0 ± 0.0	5.0 ± 0.0	8.0 ± 0.0	C
ความละเอียดวิดีโอเฉลี่ย (Width)	1198.4 ± 137.1	1193.6 ± 134.8	774.4 ± 91.5	A
ความละเอียดวิดีโอเฉลี่ย (Height)	779.2 ± 101.5	774.4 ± 100.7	518.4 ± 65.4	A
Total Bitrate เฉลี่ย	2775.6 ± 1472.0	2765.2 ± 1943.5	2778.3 ± 1668.0	C
Frame Rate เฉลี่ย	24.0 ± 0.0	30.0 ± 0.0	30.0 ± 0.0	B / C
ขนาดไฟล์เฉลี่ย	1.87 ± 0.07	0.88 ± 0.04	1.06 ± 0.05	B
Size / Duration	0.374	0.176	0.133	C

ตารางที่ 3 ประสิทธิภาพด้านคุณภาพ

ตัวชี้วัด	Runway	Wan2.1	Framepack
ความละเอียดเฉลี่ย	สูง (≈1200×780)	สูง (≈1190×770)	ปานกลาง (≈770×520)
Total Bitrate	ปานกลาง	ปานกลาง	สูงเล็กน้อย
คะแนนคุณภาพเฉลี่ย (1–10)	8.0	7.5	7.8

เพื่อให้เห็นถึงความเหมาะสมในการนำไปใช้งานในสถานการณ์จริงผู้วิจัยจึงได้ทำการเปรียบเทียบระยะเวลาที่ใช้ในการประมวลผลและแหล่งทรัพยากรที่แต่ละระบบเลือกใช้

ตารางที่ 4 ประสิทธิภาพด้านการประมวลผล

ตัวชี้วัด	Runway	Wan2.1	Framepack
ความยาววิดีโอ (s)	5	5	8
เวลาในการประมวลผล (นาที)	≈ 2	≈ 5–6	≈ 45–55
ค่าใช้จ่าย	มี	มี	ไม่มี
แหล่งทรัพยากร	Cloud	Cloud	Local
ประสิทธิภาพโดยรวม	เร็ว	ปานกลาง	ช้า

จากตารางที่ 4 การทดสอบพบว่า Framepack (Model C) เป็นการประมวลผลผ่านทรัพยากรภายในเครื่องแม้จะไม่มีค่าใช้จ่ายในการใช้งานแต่ต้องใช้ระยะเวลาในการประมวลผลนานที่สุด

ตารางที่ 5 แสดงประสิทธิภาพการใช้พื้นที่จัดเก็บของแต่ละโมเดล

ตัวชี้วัด	Runway	Wan2.1	Framepack
ขนาดไฟล์เฉลี่ย (MB)	1.87	0.88	1.06
Size / Duration (MB/s)	0.374	0.176	0.133
อัตราส่วนการบีบอัด	ต่ำ	สูง	สูงมาก

เพื่อให้เห็นภาพรวมของความแตกต่างในแต่ละมิติ ผู้วิจัยได้ทำการสรุปผลการเปรียบเทียบประสิทธิภาพของระบบปัญญาประดิษฐ์ทั้ง 3 รูปแบบ ดังแสดงในตารางที่ 6

ตารางที่ 6 สรุปการเปรียบเทียบประสิทธิภาพโดยรวม

ด้านการประเมิน	Runway	Wan2.1	Framepack
คุณภาพวิดีโอ	สูง	สูง	ปานกลาง
คุณสมบัติทางเทคนิค	สูง	สูง	ปานกลาง
ความเร็วในการประมวลผล	เร็ว	ปานกลาง	ช้ามาก
การใช้พื้นที่จัดเก็บ	ต่ำ	สูง	สูงมาก
ความเหมาะสมในการใช้งาน	งานเร่งด่วน /คุณภาพสูง	งานทั่วไป / คุ้มค่า	งานไม่เร่ง

จากตารางที่ 6 Runway (Model A) เหมาะกับงานเร่งด่วนที่ต้องการคุณภาพสูงแต่มีค่าบริการ Wan2.1 (Model B) เหมาะกับงานทั่วไปที่เน้นความคุ้มค่า และ Framepack (Model C) เหมาะกับงานที่ไม่เร่งด่วนและต้องการประหยัดค่าใช้จ่าย แต่ใช้เวลาประมวลผลนานกว่า

ผู้วิจัยยังทดสอบเปรียบเทียบระบบสังเคราะห์เสียงภาษาไทย 2 โมเดล ได้แก่ Botnoi และ F5-TTS-Thai โดยใช้ 10 ประโยคทดสอบที่หลากหลายทั้งความยาว ความซับซ้อน และคำศัพท์เทคนิค เพื่อวิเคราะห์ด้านเทคนิคและคุณภาพการออกเสียง ดังนี้

ตารางที่ 7 ชุดประโยคทดสอบ

ตัวอย่างที่	ประโยคทดสอบ	ความยาว (คำ)
S01	ในอวกาศอันกว้างใหญ่ มีหินและเศษซากล่องลอยอยู่มากมาย ส่วนใหญ่อยู่รวมกันในแถบดาวเคราะห์น้อย แต่บางครั้งก็แตกกระจายเพราะการปะทะ	26
S02	เมื่อดาวเคราะห์น้อยชนกันด้วยความเร็วสูง เศษหินและฝุ่นถูกเหวี่ยงออกไปอย่างรุนแรง และบางส่วนก็หลุดออกจากวงโคจรเดิม	24
S03	เศษหินที่หลุดออกมากลายเป็นอุกกาบาต พวกมันโคจรอิสระและอาจพุ่งสู่ชั้นบรรยากาศโลก แสงวาบยามค่ำคืนก็คือดาวตก	23
S04	หากอุกกาบาตมีขนาดใหญ่ มันจะไม่เผาไหม้หมดในชั้นบรรยากาศแต่พุ่งชนพื้นโลก เกิดเป็นหลุมมหึมาอย่างทรงพลัง	22
S05	ระบบสุริยะประกอบด้วยดาวเคราะห์แปดดวง โคจรรอบดวงอาทิตย์ โดยมีอายุประมาณ 4.6 พันล้านปี	17
S06	กล้องโทรทรรศน์อวกาศเจมส์เวบบ์สามารถสังเกตดาวเคราะห์นอกระบบสุริยะที่อยู่ห่างออกไปหลายปีแสง	17
S07	ซูเปอร์โนวาคือการระเบิดของดาวฤกษ์ขนาดใหญ่ ปลดปล่อยพลังงานมหาศาลภายในเวลาอันสั้น	15
S08	ทางช้างเผือกเป็นกาแล็กซีแบบกังหัน มีดาวฤกษ์มากกว่า 100 พันล้านดวง	12
S09	การศึกษาคลื่นความโน้มถ่วงช่วยยืนยันการรวมตัวของหลุมดำสองแห่งในเอกภพอันกว้างใหญ่	16
S10	ดาวฤกษ์เกิดจากการยุบตัวของกลุ่มก๊าซไฮโดรเจนภายใต้แรงโน้มถ่วงเป็นเวลาหลายล้านปี	16

จากประโยคทดสอบ 10 ชุดในตารางที่ 7 ผู้วิจัยประเมินคุณภาพเสียงด้วยเกณฑ์คะแนน 1–5 โดยผลคะแนนเฉลี่ยแสดงในตารางที่ 8

ตารางที่ 8 ค่าเฉลี่ยคะแนนการประเมินคุณภาพเสียง

ด้านการประเมินคุณภาพ	Botnoi	F5-TTS-Thai
ความเข้าใจ (intelligibility)	**4.89**	4.56
ความเป็นธรรมชาติ (naturalness)	4.10	**4.50**
การออกเสียง (pronunciation)	3.80	**4.11**
จังหวะและน้ำเสียง (prosody)	**4.20**	3.89
คุณภาพเสียง (audio quality)	**5.00**	**5.00**
คะแนนเฉลี่ยรวม	**4.40**	**4.41**

เพื่อเปรียบเทียบความแตกต่างในด้านอื่นนอกเหนือจากคุณภาพเสียง ผู้วิจัยสรุปผลการเปรียบเทียบดังตารางที่ 9

ตารางที่ 9 สรุปการเปรียบเทียบประสิทธิภาพคุณภาพและค่าใช้จ่าย

ด้านการประเมิน	Botnoi	F5-TTS-Thai
จุดเด่นเชิงคุณภาพ	ความชัดเจนและจังหวะเสียง	เสียงธรรมชาติ ออกเสียงแม่นยำ
รูปแบบการประมวลผล	คลาวด์ (สะดวกและรวดเร็ว)	ประมวลผลภายในเครื่อง
ค่าใช้จ่าย	**มีค่าใช้จ่าย**	**ไม่มีค่าใช้จ่าย**
ความเหมาะสมในการใช้งาน	งานเร่งด่วน / ชัดเจน	งานธรรมชาติ / ประหยัดงบ

จากตารางที่ 9 F5-TTS-Thai เป็นทางเลือกที่คุ้มค่าที่สุด เนื่องจากให้เสียงเป็นธรรมชาติและไม่มีค่าใช้จ่ายระยะยาว ส่วน Botnoi เหมาะกับงานที่ต้องการความสะดวกและรวดเร็วในการติดตั้ง

5. สรุปผลและอภิปรายผล

จากการทดลองพบว่า Runway (Model A) มีประสิทธิภาพสูงสุดด้านคุณภาพและความเร็ว โดยให้ความละเอียดเฉลี่ย 1,200×780 พิกเซล คะแนนความพึงพอใจร้อยละ 80 และใช้เวลาประมวลผลเพียง 2 นาที รองลงมาคือ Framepack (Model C) ได้คะแนนร้อยละ 78 สร้างวิดีโอได้ยาวสูงสุด 8.0 วินาที แต่ใช้เวลาประมวลผล 45–55 นาที ส่วน Wan2.1 (Model B) ได้คะแนนร้อยละ 75 และมีขนาดไฟล์เฉลี่ยต่ำสุดที่ 0.88 MB สำหรับระบบสังเคราะห์เสียงภาษาไทย F5-TTS-Thai และ Botnoi มีคะแนนรวมใกล้เคียงกันที่ 4.41 และ 4.40 ตามลำดับ โดย F5-TTS-Thai มีความโดดเด่นด้านความเป็นธรรมชาติของเสียงและไม่มีค่าใช้จ่าย ขณะที่ Botnoi เด่นด้านความเข้าใจและจังหวะน้ำเสียง

ผลการทดลองแสดงให้เห็นความสัมพันธ์เชิงแลกเปลี่ยนระหว่างคุณภาพ ความเร็ว และต้นทุน โดย Runway เหมาะกับงานที่ต้องการคุณภาพและความเร็วสูงแต่มีค่าใช้จ่ายตามปริมาณการใช้งาน ขณะที่ Framepack และ F5-TTS-Thai สามารถให้ผลลัพธ์ในระดับที่ยอมรับได้โดยไม่มีค่าซอฟต์แวร์ อย่างไรก็ตาม ยังมีข้อจำกัดที่ควรพิจารณา ได้แก่ ความต้องการทรัพยากร GPU ประสิทธิภาพสูง ความปลอดภัยของข้อมูลในระบบคลาวด์ และความแม่นยำในการประมวลผลคำศัพท์เฉพาะทาง

สำหรับแนวทางการพัฒนาต่อยอดผู้วิจัยเสนอ 3 ประการ ได้แก่ การอัปเกรดโมเดลตามความเหมาะสมของงบประมาณ การเชื่อมโยงเฟรมภาพระหว่างฉากเพื่อความต่อเนื่องของเนื้อหา และการพัฒนาทักษะด้านการออกแบบเวิร์กโฟลว์อัตโนมัติ

6. กิตติกรรมประกาศ

ขอขอบคุณสาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้และศูนย์ปฏิบัติการหอดูดาวงานวิศวกรรมเมคาทรอนิกส์ สถาบันวิจัยดาราศาสตร์แห่งชาติ (องค์การมหาชน) ที่ให้การสนับสนุนสถานที่และข้อมูลสำหรับการวิจัยในครั้งนี้

7. เอกสารอ้างอิง

ผุสดี บุญรอด, & ประกายมาศ ศรีสุขทักษิณ. (2015). การค้นคืนข้อมูลขนาดใหญ่โดยใช้ภาษสอบถามแบบไม่มีโครงสร้างร่วมกับเทคโนโลยีเว็บเชิงความหมาย. *วารสารวิชาการพระจอมเกล้าพระนครเหนือ, 25*(2), 255-264. https://ph01.tci-thaijo.org/index.php/kmutnb-journal/article/view/33784

Adobe. (n.d.). *คำแนะนำสำหรับมือใหม่เกี่ยวกับความละเอียดวิดีโอ.* Retrieved February 28, 2026, from https://www.adobe.com/th_th/creativecloud/video/discover/video-resolution.html

Hewlett Packard Enterprise. (2025). *What are computing resources?.* Retrieved March 17, 2026, from https://www.hpe.com/emea_europe/en/what-is/compute-resources.htm

Kaur, N., & Singh, P. (2023). Conventional and contemporary approaches used in text to speech synthesis: A review. *Artificial Intelligence Review, 56*(7), 5837-5880. https://doi.org/10.1007/s10462-022-10315-0

Venkiteela, P. (2025). n8n: An Open-Source Workflow Automation Platform for Enterprise Integration and AI-Driven Orchestration. *International Journal of Computer Applications*, *187*(63), 1–11. https://doi.org/10.5120/ijca2025926031

Xie, W., Hu, A., Xie, Q., Chen, J., Wan, R., & Liu, Y. (2025). Bibliometric analysis and review of AI-based video generation: research dynamics and application trends (2020–2025). *Discover Computing, 28*(1), 130. https://doi.org/10.1007/s10791-025-09628-9

Yan, L., Han, C., Xu, Z., Liu, D., & Wang, Q. (2023, August). Prompt learns prompt: Exploring knowledge-aware generative prompt collaboration for video captioning. *In IJCAI* (pp. 1622-1630).