การพัฒนาระบบแชทบอทสำหรับเว็บไซต์มหาวิทยาลัยโดยใช้โมเดลภาษาขนาดใหญ่และการสร้างคำตอบด้วยการค้นคืนผ่านแพลตฟอร์ม Flowise

กฤษฎิ์ภวิศ สูงสว่าง[1] ธนพร ปฏิกรณ์[2*] เอกชัย เนาวนิช[3]

Received : May 6, 2025
Revised : December 19, 2025
Accepted : December 26, 2025

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบแชทบอทสำหรับเว็บไซต์มหาวิทยาลัย โดยใช้โมเดลภาษาขนาดใหญ่ (Large Language Models: LLMs) และเทคนิคการสร้างคำตอบด้วยการค้นคืน (Retrieval-Augmented Generation: RAG) ผ่านแพลตฟอร์ม Flowise ที่ช่วยลดความซับซ้อนในการพัฒนาแชทบอทด้วยแนวคิด low-code/no-code ข้อมูลที่ใช้พัฒนามาจากเว็บไซต์มหาวิทยาลัยและหลักสูตรต่าง ๆ โดยจัดเก็บในรูปแบบไฟล์ข้อความและเวกเตอร์ในฐานข้อมูล Postgres (pgvector) พร้อมบันทึกประวัติการสนทนาในฐานข้อมูล MongoDB กระบวนการพัฒนาประกอบด้วยการเตรียมข้อมูล การตั้งค่าโมเดลเอมเบดดิง (Embedding) และคลังเอกสาร (Document Store) แล้วทำการออกแบบ Chatflows เพื่อสร้างแชทบอทที่ตอบคำถามเกี่ยวกับมหาวิทยาลัยและหลักสูตรได้ ระบบได้รับการทดสอบและประเมินผลเชิงคุณภาพ พบว่าคำตอบส่วนใหญ่ถูกต้องและเหมาะสม แต่บางคำตอบที่ไม่ครบถ้วนหรือผิดพลาดอาจเกิดจากการจัดการข้อมูลและการตั้งค่าที่ไม่เหมาะสม

ผลการประเมินผลการทำงานของแชทบอทในภาพรวม พบว่า แชทบอทสามารถตอบคำถามส่วนใหญ่ได้ถูกต้อง เช่น คำถามเกี่ยวกับประวัติมหาวิทยาลัย การบริหารงาน และสถานที่ติดต่อ แต่สำหรับบางคำถามที่แชทบอทไม่สามารถให้ข้อมูลได้ถูกต้องและครบถ้วน เช่น การขอทราบรายการหลักสูตรต่าง ๆ ในคณะ การวิจัยนี้แสดงให้เห็นถึงศักยภาพของแพลตฟอร์ม Flowise ในการพัฒนาแชทบอทที่ลดระยะเวลาและความซับซ้อน พร้อมทั้งเป็นแนวทางสำหรับการนำเทคโนโลยี AI ไปประยุกต์ใช้ในสถาบันการศึกษาในอนาคต

คำสำคัญ: การสร้างคำตอบด้วยการค้นคืน แชทบอท แพลตฟอร์ม low-code/no-code โมเดลภาษาขนาดใหญ่ Flowise

[1] สาขาวิชาเทคโนโลยีดิจิทัลมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ อีเมล: 167491432001-st@rmutsb.ac.th

[2] สาขาวิชาเทคโนโลยีสารสนเทศและเศรษฐกิจดิจิทัล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ อีเมล: thanaporn.pa@rmutsb.ac.th

[3] สาขาวิชาเทคโนโลยีดิจิทัลมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ อีเมล: ekachai.n@rmutsb.ac.th

[*] ผู้นิพนธ์หลัก อีเมล: thanaporn.pa@rmutsb.ac.th

DEVELOPING A CHATBOT SYSTEM FOR UNIVERSITY WEBSITE USING LARGE LANGUAGE MODELS AND RETRIEVAL-AUGMENTED GENERATION VIA FLOWISE PLATFORM

Kritpawit Soongswang[1] Thanaporn Patikorn[2*] Ekachai Naowanich[3]

Abstract

This research aimed to develop a chatbot system for a university website using Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) techniques via the Flowise platform, which simplifies chatbot development through low-code/no-code concepts. The data utilized for development were obtained from the university website and various course catalogs, stored as text files and vector embeddings in a Postgres database with pgvector extension, while conversation histories were recorded in MongoDB. The development process comprised data preparation, embedding model configuration, document store setup, and chatflow design to create a chatbot capable of answering queries related to the university and its academic programs. The system underwent qualitative testing and evaluation, revealing that most responses were accurate and appropriate. However, some incomplete or inaccurate responses were attributed to suboptimal data management and configuration settings.

Overall performance evaluation indicated that the chatbot successfully answered most inquiries, particularly those concerning university history, administrative structure, and contact information. Nevertheless, the system encountered difficulties in providing complete and accurate information for certain queries, such as comprehensive lists of academic programs within faculties. This study demonstrates the potential of the Flowise platform in streamlining chatbot development by reducing both time requirements and technical complexity, while also providing valuable guidance for implementing AI technologies in educational institutions.

Keywords: Retrieval-augmented generation, Chatbot, Low-code/no-code platform, Large language models, Flowise

---

[1] *Digital Media Technology Program, Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi, e-mail: 167491432001-st@rmutsb.ac.th*

[2] *Information Technology and Digital Economy Program, Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi, e-mail: thanaporn.pa@rmutsb.ac.th*

[3] *Digital Media Technology Program, Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi, e-mail: ekachai.n@rmutsb.ac.th*

[*] *Corresponding author, e-mail: thanaporn.pa@rmutsb.ac.th*

บทนำ

ในยุคดิจิทัลที่เทคโนโลยีมีการเปลี่ยนแปลงอย่างรวดเร็ว การพัฒนาระบบที่สามารถตอบสนองความต้องการของผู้ใช้งานได้อย่างมีประสิทธิภาพและทันท่วงทีเป็นสิ่งจำเป็นอย่างยิ่ง โดยเฉพาะในบริบทของสถาบันการศึกษาระดับอุดมศึกษาที่ต้องการนำเสนอข้อมูลที่หลากหลายและซับซ้อน เช่น หลักสูตรการเรียน การรับสมัครนักศึกษา ทุนการศึกษา และกิจกรรมต่าง ๆ ซึ่งในอดีตมักอยู่ในรูปแบบที่ไม่เปลี่ยนแปลง (static) เช่น คำถามที่มักจะถามบ่อย (FAQ) และไฟล์เอกสาร ซึ่งมันมีข้อมูลมหาศาลและค้นหาได้ยาก หรือการใช้เจ้าหน้าที่คอยตอบคำถาม ซึ่งใช้ทรัพยากรบุคคลและเวลามาก ทำให้ในปัจจุบันระบบแชทบอทได้รับการยอมรับอย่างกว้างขวางว่าเป็นเทคโนโลยีที่มีศักยภาพในการตอบสนองความต้องการเหล่านี้ (Aloqayli & Abdelhafez, 2023) เนื่องจากสามารถให้บริการข้อมูลได้อย่างรวดเร็ว สะดวก และมีความพร้อมใช้งานตลอด 24 ชั่วโมง (Radziwill & Benton, 2017)

การพัฒนาของเทคโนโลยีปัญญาประดิษฐ์ โดยเฉพาะแบบจำลองภาษาขนาดใหญ่ (Large Language Models: LLMs) ได้สร้างการเปลี่ยนแปลงครั้งสำคัญในวงการประมวลผลภาษาธรรมชาติ (Brown et al., 2020) เมื่อนำมาผนวกกับเทคนิคการสร้างคำตอบด้วยการค้นคืน (Retrieval-Augmented Generation: RAG) ทำให้ระบบแชทบอทมีความสามารถในการเข้าใจบริบท วิเคราะห์คำถาม และดึงข้อมูลจากแหล่งข้อมูลที่กำหนดได้อย่างแม่นยำมากขึ้น (Lewis et al., 2020) นอกจากนี้ การใช้ RAG ยังช่วยลดปัญหา Hallucination หรือการสร้างข้อมูลที่ไม่ถูกต้อง ซึ่งเป็นข้อจำกัดสำคัญของ LLMs ทั่วไป (Izacard & Grave, 2021)

อย่างไรก็ตาม การพัฒนาระบบแชทบอทที่ใช้โมเดลภาษาขนาดใหญ่ (LLMs) และเทคนิคการสร้างคำตอบด้วยการค้นคืนข้อมูล (RAG) นั้น มักเป็นกระบวนการที่ซับซ้อนและต้องการความเชี่ยวชาญทางเทคนิคสูง ผู้พัฒนาต้องเผชิญกับความท้าทายในการตั้งค่าโมเดล การจัดการข้อมูล (Arslan et al, 2024) และการเขียนโค้ดที่ซับซ้อนเพื่อให้ระบบสามารถทำงานได้ตามที่ต้องการ (Yu et al, 2024) ด้วยเหตุนี้ แนวคิด low-code/no-code platforms จึงถูกพัฒนาขึ้นเพื่อช่วยลดความซับซ้อนเหล่านี้ โดยเน้นให้ผู้ใช้งานสามารถสร้างแอปพลิเคชัน AI ได้ผ่านอินเทอร์เฟซที่ไม่ต้องพึ่งพาการเขียนโค้ดในปริมาณมาก (Tunstall et al., 2022)

Flowise เป็นตัวอย่างหนึ่งของแพลตฟอร์ม low-code/no-code ที่ถูกออกแบบมาเพื่อตอบสนองความต้องการในการพัฒนาระบบแชทบอท AI ด้วยการเชื่อมต่อกับ LLMs และกระบวนการค้นคืนข้อมูล (Izacard, 2025) โดยที่ Flowise ช่วยให้องค์กรสามารถสร้างแชทบอทที่มีความสามารถในการตอบสนองคำถามที่ซับซ้อนได้ โดยไม่จำเป็นต้องมีทีมพัฒนาที่มีความเชี่ยวชาญเฉพาะทางด้าน AI (Reis et al., 2024) สิ่งนี้ช่วยลดภาระในกระบวนการเขียนโค้ดที่ซับซ้อน และเพิ่มความรวดเร็วในการพัฒนาระบบ ทำให้องค์กรสามารถนำเทคโนโลยี AI มาประยุกต์ใช้ในบริบทต่าง ๆ ได้อย่างมีประสิทธิภาพ ซึ่งในปัจจุบันยังไม่มีงานวิจัยที่ตีพิมพ์ที่นำ Flowise มาประยุกต์ใช้ในบริบทมหาวิทยาลัยไทย

งานวิจัยนี้ ระบบแชทบอทจะถูกพัฒนาขึ้นโดยใช้ชุดข้อมูลที่เกี่ยวข้องกับข้อมูลมหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ เช่น ข้อมูลเว็บไซต์มหาวิทยาลัย ที่ครอบคลุมข้อมูลทั่วไป ได้แก่ ประวัติของมหาวิทยาลัย วิสัยทัศน์ พันธกิจ โครงสร้างการบริหาร ค่านิยม หลักสูตร และข้อมูลสำหรับผู้สนใจศึกษาต่อ รวมถึงข้อมูลเฉพาะด้าน เช่น ข้อมูลหลักสูตรของคณะวิทยาศาสตร์และเทคโนโลยี ที่ประกอบด้วยหลักสูตรในศูนย์การศึกษาต่าง ๆ เช่น นนทบุรี สุพรรณบุรี และพระนครศรีอยุธยา ข้อมูลดังกล่าวถูกจัดเตรียมในรูปแบบไฟล์ข้อความที่พร้อมใช้งานในระบบ Flowise

งานวิจัยนี้มุ่งศึกษาและพัฒนาระบบแชทบอทสำหรับเว็บไซต์มหาวิทยาลัย โดยใช้ Flowise เป็นเครื่องมือหลักในการพัฒนา เพื่อสร้างระบบที่สามารถให้ข้อมูลเกี่ยวกับมหาวิทยาลัย ได้อย่างมีประสิทธิภาพ การศึกษานี้จะช่วยสร้างองค์ความรู้และแนวทางปฏิบัติสำหรับสถาบันการศึกษาที่ต้องการพัฒนาระบบแชทบอทในอนาคต รวมถึงแสดงให้เห็นถึงศักยภาพของการนำเทคโนโลยี LLMs และ RAG มาประยุกต์ใช้ในบริบทของการศึกษาระดับอุดมศึกษา

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาระบบแชทบอทสำหรับเว็บไซต์มหาวิทยาลัย โดยใช้ Flowise เป็นเครื่องมือหลักในการพัฒนา โดยเน้นการให้ข้อมูลเกี่ยวกับมหาวิทยาลัย และหลักสูตร

2. เพื่อสร้างกระบวนการพัฒนาระบบแชทบอทที่เชื่อมต่อกับ LLMs และ RAG โดยลดความซับซ้อนผ่านแพลตฟอร์ม low-code/no-code

3. เพื่อประเมินความเหมาะสมและความถูกต้องของคำตอบที่ระบบแชทบอทตอบ โดยการตรวจสอบความสอดคล้องกับข้อมูลต้นฉบับ พร้อมทั้งระบุข้อดีและข้อจำกัดเพื่อเป็นแนวทางในการพัฒนาระบบในอนาคต

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยและพัฒนา แชทบอท สำหรับเว็บไซต์มหาวิทยาลัย ด้วย Flowise เป็นเครื่องมือหลักในการพัฒนา ซึ่งมีขั้นตอนวิธีดำเนินการวิจัย ดังต่อไปนี้

ขั้นตอนที่ 1 การเตรียมสภาพแวดล้อม สำหรับ การติดตั้งแพลตฟอร์ม Flowise

ขั้นตอนนี้จะดำเนินการเตรียมความพร้อมของฮาร์ดแวร์และซอฟต์แวร์ที่จำเป็นสำหรับการติดตั้งระบบ Flowise เพื่อใช้สำหรับเป็นเครื่องมือในการพัฒนา ระบบแชทบอท ดังนี้

1. อุปกรณ์ที่ใช้ในการทดลอง

งานวิจัยนี้ถูกพัฒนาบนแลปท็อปที่มีคุณสมบัติ ได้แก่ หน่วยประมวลผล Intel® Core™ i7-1360P หน่วยความจำระยะสั้น (RAM) ขนาด 32 GB และหน่วยความจำระยะยาว (Storage) แบบ SSD ความจุ

951 GB โดยทำงานบนระบบปฏิบัติการ Windows 11 Pro และสามารถเชื่อมต่อกับอินเทอร์เน็ตความเร็วสูง เพื่อรองรับการเชื่อมต่อ Application Programming Interface (API) ของ LLMs ได้อย่างมีประสิทธิภาพ

2. ซอฟต์แวร์สำหรับการติดตั้ง และเริ่มต้นแพลตฟอร์ม Flowise

งานวิจัยนี้ใช้ Node.js เวอร์ชั่น 20.12.2 ซึ่งเป็นเครื่องมือสำคัญสำหรับการ เริ่มต้นระบบ Flowise (https://flowiseai.com/) และจัดการโมดูล Javascript

3. ระบบฐานข้อมูลที่ใช้ในงานวิจัย

งานวิจัยนี้ ระบบฐานข้อมูลได้ถูกแบ่งออกเป็น 2 ส่วนหลัก เพื่อรองรับการทำงานของแชทบอท ได้แก่

3.1 ฐานข้อมูลสำหรับการค้นคืนเอกสาร (ฐานข้อมูลเวกเตอร์)

3.1.1 ใช้ฐานข้อมูลที่ชื่อว่า pgvector ซึ่งเป็น ส่วนขยายของ PostgreSQL ในการจัดเก็บข้อมูลเวกเตอร์ ทำให้การดึงข้อมูลที่มีบริบทใกล้เคียงเป็นไปอย่างรวดเร็วและแม่นยำ

3.1.2 ข้อมูลที่จัดเก็บในฐานข้อมูลนี้ ได้แก่ ข้อมูลจากเว็บไซต์มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิและข้อมูลหลักสูตรต่าง ๆ

3.2 ฐานข้อมูลสำหรับการเก็บประวัติการสนทนา

3.2.1 ใช้ฐานข้อมูลที่ชื่อว่า MongoDB ซึ่งเป็นฐานข้อมูลแบบ NoSQL ที่มีความยืดหยุ่นและมีความเร็วในการอ่านและเขียนข้อมูลสูง

โดยทั้งสองฐานข้อมูลนี้ ถูกติดตั้งและจัดการผ่าน Docker ซึ่งเป็นแพลตฟอร์มที่ช่วยสร้างและจัดการคอนเทนเนอร์ซอฟต์แวร์ ทำให้การติดตั้งฐานข้อมูลเป็นไปอย่างสะดวกและสามารถควบคุมสภาพแวดล้อมได้อย่างมีประสิทธิภาพ

ขั้นตอนที่ 2 การเตรียมและทำความสะอาดข้อมูล

ข้อมูลที่ใช้ในงานวิจัยนี้แบ่งออกเป็น 2 ส่วนหลัก ได้แก่ ข้อมูลจากเว็บไซต์มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ (https://rmutsb.ac.th/home) และ ข้อมูลหลักสูตรต่าง ๆ ในเว็บไซต์ของคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ (https://sci.rmutsb.ac.th/) โดยการทำความสะอาดข้อมูล สามารถแบ่งประเภทของข้อมูลจากหน้าเว็บไซต์ เป็น 3 ประเภท ได้แก่ ข้อมูลในเว็บไซต์ที่ไม่ซับซ้อน ข้อมูลที่เป็นรูปภาพแผนผัง และข้อมูลในเว็บไซต์ที่มีความซับซ้อน

1. เว็บไซต์ที่ไม่ซับซ้อน คือเว็บไซต์ที่ข้อมูลถูกจัดเรียงอย่างเป็นระเบียบ และสามารถคัดลอกข้อมูลได้ทั้งหมดในครั้งเดียว โดยสามารถทำความสะอาดข้อมูลได้ ดังนี้

1.1 คัดลอกข้อความจากหน้าเว็บไซต์มหาวิทยาลัย

1.2 ใช้ความสามารถของ ChatGPT-4o จาก OpenAI ในการจัดระเบียบและปรับโครงสร้างข้อมูล โดยใช้คำสั่งว่า “พิมและจัดเรียงข้อมูลทั้งหมด โดยที่ห้ามตัดเนื้อหาออก จงใช้เนื้อหาเดิมเท่านั้น ห้ามสร้างเนื้อหาใหม่ โดยตอบแค่ผลลัพธ์เท่านั้น”

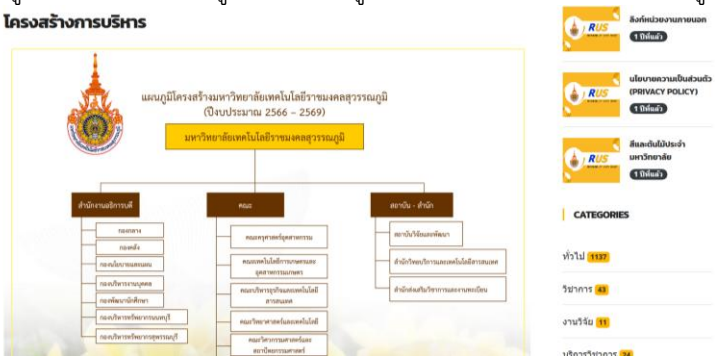
1.3 ผู้วิจัยตรวจสอบความถูกต้องของข้อมูลที่สกัดได้จาก LLM และบันทึกในรูปแบบ text file

ภาพที่ 1 ตัวอย่างเว็บไซต์ที่แสดงข้อมูลที่ไม่ซับซ้อน

2. การทำความสะอาดข้อมูลจากเว็บไซต์ ที่ให้ข้อมูลเป็นรูปภาพ แผนผัง ดังนี้

2.1 ใช้ Vision Model ใน ChatGPT-4o จาก OpenAI เพื่ออ่านและสกัดข้อมูลจากแผนภาพและอินโฟกราฟิก ซึ่งต่างกับการสกัดข้อมูลจากหน้าเว็บไซต์ปกติ เนื่องจาก Vision Model นั้นสามารถสกัดข้อมูลจากภาพและไฟล์เอกสารสแกนได้ด้วย

2.2 ผู้วิจัยตรวจสอบความถูกต้องของข้อมูลที่สกัดได้จาก LLM และบันทึกในรูปแบบ text file



ภาพที่ 2 ตัวอย่างเว็บไซต์ที่แสดงข้อมูลรูปภาพแผนผัง

3. เว็บไซต์ที่มีความซับซ้อน คือเว็บไซต์ที่ข้อมูลไม่ได้ถูกวางอย่างเป็นระเบียบ และมีส่วนของข้อมูลที่ไม่สามารถคัดลอกได้เนื่องจากการแสดงผลที่ซับซ้อน เช่น ต้องมีการคลิกที่หัวข้อเพื่อเปิดอ่านข้อความภายในดังภาพที่ 3 โดยสามารถทำความสะอาดข้อมูลได้ ดังนี้

3.1 ใช้การ inspect HTML element จากหน้าเว็บไซต์เพื่อระบุข้อมูลที่ต้องการ จากนั้นคัดลอกบล็อคข้อมูลที่สำคัญ โดยใช้คำสั่งว่า “สกัดข้อมูล จาก html element โดยรักษาเนื้อหาทั้งหมด โดยตอบแค่ผลลัพธ์เท่านั้น”

3.2 ใช้ความสามารถของ ChatGPT-4o ในการสกัดข้อมูลจาก HTML element และกรณีที่มีขนาดข้อมูลใหญ่มาก จะใช้ ChatGPT o1-preview แทน เนื่องจาก ChatGPT-4o นั้นมีความจำกัดทางด้านขนาดของอินพุทที่ใช้ในการประมวลผล

3.3 ผู้วิจัยตรวจสอบความถูกต้องของข้อมูลที่สกัดได้จาก LLM และบันทึกในรูปแบบ text file

ภาพที่ 3 ตัวอย่างเว็บไซต์ที่มีความซับซ้อน

โดยหลังจากทำความสะอาดข้อมูล สามารถรวบรวมข้อมูลได้จำนวนทั้งสิ้น 37 ไฟล์ ซึ่งรวบรวมมาจากเว็บไซต์จำนวน 36 หน้า โดยแบ่งออกเป็น เว็บไซต์มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิจำนวน 22 ไฟล์ และเว็บไซต์ของคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิจำนวน 15 ไฟล์

ขั้นตอนที่ 3 การพัฒนาระบบแชทบอต ผ่านแพลตฟอร์ม Flowise

1. ทำการสร้าง Document Store เพื่อจัดเก็บไฟล์เอกสารต่าง ๆ ที่อยู่ในรูปของ text file

1.1 เลือก การตัดคำเป็น Recursive Character Text Splitter เพื่อทำการแบ่งข้อความ เป็นส่วน ๆ เพื่อให้ Embedding Model สามารถแปลงข้อความที่ไม่ยาวเกินขีดจำกัดของข้อความที่โมเดลรับได้ ให้อยู่ในรูปแบบของเวกเตอร์เพื่อให้รองรับกับการสืบค้นเอกสาร

1.2 กำหนดความยาวในแต่ละส่วน เป็น 8000 คำ (Chunk Size) เพื่อรักษาความต่อเนื่องของข้อมูลในแต่ละไฟล์

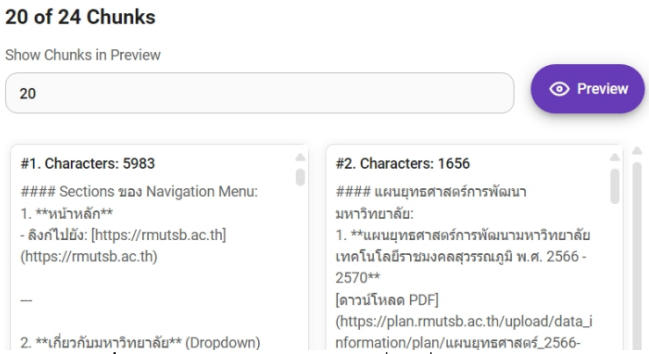
1.3 กำหนดช่วงทับซ้อนระหว่าง ก้อนข้อความ (Chunk Overlap) เป็น 500 คำ เพื่อรักษาความต่อเนื่องของข้อมูลในแต่ละไฟล์ที่ถูกแบ่งออกเป็นหลายก้อน กรณีที่ไฟล์นั้นมีข้อความยาวเกิน 8000 คำ

1.4 อัพโหลดไฟล์ และกระบวนการตัดและแบ่งคำออกเป็นกลุ่มก้อน โดยหลังจากเสร็จกระบวนการจะได้ก้อนข้อมูลจำนวนทั้งสิ้น 39 ก้อน

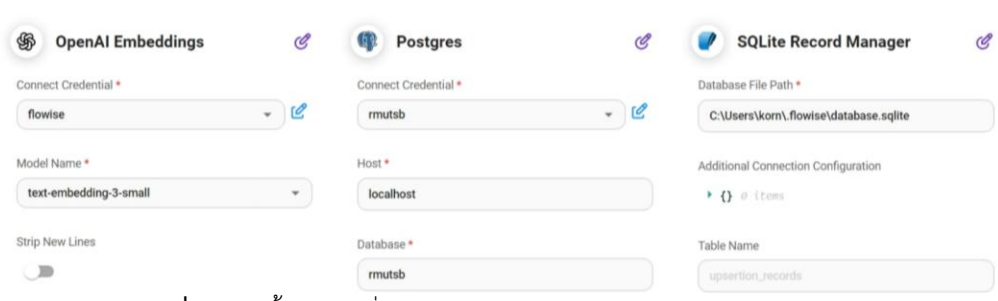
ภาพที่ 4 หน้าจอ Document Store

2. ทำการตั้งค่าการเชื่อมต่อกับ Embedding Model ซึ่งเป็นโมเดลปัญญาประดิษฐ์ ที่ทำหน้าที่แปลงก้อนข้อความ หรือคำถามของผู้ใช้เป็นเวกเตอร์เพื่อเก็บข้อมูล โดยใช้โมเดล text-embedding-3-small จาก OpenAI เนื่องจากคิดค่าบริการที่ไม่สูง

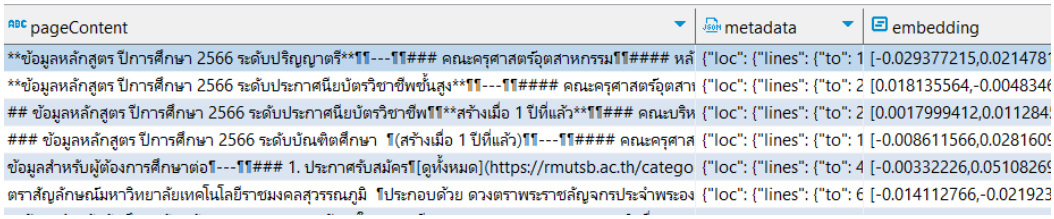
3. ทำการตั้งค่าการเชื่อมต่อกับฐานข้อมูลเวกเตอร์ เพื่อจัดเก็บข้อมูลของก้อนข้อความ พร้อมเวกเตอร์ของแต่ละก้อนข้อความลงบนฐานข้อมูลโดยทำการเลือกฐานข้อมูล Postgres และเลือกไดร์ฟเวอร์เป็น pgvector จากนั้น ตั้งค่า Record Manager โดยใช้ระบบตำแหน่งของ SQLite เพื่อทำหน้าที่เป็นฐานข้อมูลสำหรับบันทึกการเปลี่ยนแปลงของ pgvector เมื่อตั้งค่าเรียบร้อยแล้ว กดปุ่ม upsert เพื่อเริ่มกระบวนการแปลงก้อนข้อมูลเป็นเวกเตอร์ และจัดเก็บเข้าสู่ฐานข้อมูล เมื่อ upsert เสร็จ ข้อมูลเอกสารและเวกเตอร์ของเอกสารจะถูกจัดเก็บไว้ในฐานข้อมูลแล้วดังภาพที่ 7 ฐานข้อมูลเวกเตอร์นั้นมีความสามารถในการรองรับการค้นหาข้อมูลโดยแปลงข้อความที่ผู้ใช้ต้องการค้นหาเป็นเวกเตอร์ผ่าน Embedding Model จากนั้นจะทำการคำนวณค่าความใกล้เคียงของข้อมูลในฐานข้อมูลโดยใช้สูตร Cosine Similarity จากนั้นฐานข้อมูลจะจัดอันดับเวกเตอร์ที่มีความใกล้เคียงกับข้อความที่ต้องการค้นหามากที่สุด และคืนลำดับของข้อความนั้น ๆ ไปยัง LLM เพื่อสร้างคำตอบไปให้ผู้ใช้งานต่อไป



ภาพที่ 5 ตัวอย่างข้อมูลจากไฟล์ที่ถูกหั่นออกเป็นก้อน



ภาพที่ 6 การตั้งค่าการเชื่อมต่อกับ Embedding Model และฐานข้อมูลเวกเตอร์



pageContent	metadata	embedding
ข้อมูลหลักสูตร ปีการศึกษา 2566 ระดับปริญญาตรี¶¶---¶¶### คณะครุศาสตร์อุตสาหกรรม¶¶#### หลั	{"loc": {"lines": {"to": 1	[-0.029377215,0.021478
ข้อมูลหลักสูตร ปีการศึกษา 2566 ระดับประกาศนียบัตรวิชาชีพชั้นสูง¶¶---¶¶#### คณะครุศาสตร์อุตสาห	{"loc": {"lines": {"to": 2	[0.018135564,-0.0048346
## ข้อมูลหลักสูตร ปีการศึกษา 2566 ระดับประกาศนียบัตรวิชาชีพ¶¶**สร้างเมื่อ 1 ปีที่แล้ว**¶¶### คณะบริห	{"loc": {"lines": {"to": 2	[0.0017999412,0.011284
### ข้อมูลหลักสูตร ปีการศึกษา 2566 ระดับบัณฑิตศึกษา ¶(สร้างเมื่อ 1 ปีที่แล้ว)¶¶---¶¶#### คณะครุศาส	{"loc": {"lines": {"to": 1	[-0.008611566,0.028160
ข้อมูลสำหรับผู้ต้องการศึกษาต่อ¶---¶¶### 1. ประกาศรับสมัคร¶[ดูทั้งหมด](https://rmutsb.ac.th/catego	{"loc": {"lines": {"to": 4	[-0.00332226,0.0510826
ตราสัญลักษณ์มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ¶ประกอบด้วย ดวงตราพระราชลัญจกรประจำพระองค์	{"loc": {"lines": {"to": 6	[-0.014112766,-0.021923

ภาพที่ 7 ก้อนข้อมูลและเวกเตอร์ ที่ถูกจัดเก็บอยู่ในฐานข้อมูล Postgresql

4. ทำการออกแบบ Chatflows ในแพลตฟอร์ม Flowise

4.1 เลือก ChatOpenAI โดยใช้โมเดล gpt-4o-mini เป็นโมเดลภาษาขนาดใหญ่ เนื่องจากเป็นโมเดลที่ได้รับการยอมรับในวงกว้าง และมีค่าบริการที่ไม่สูง

4.2 เลือก MongoDB Atlas Chat Memory เพื่อใช้ในการเก็บประวัติการสนทนา

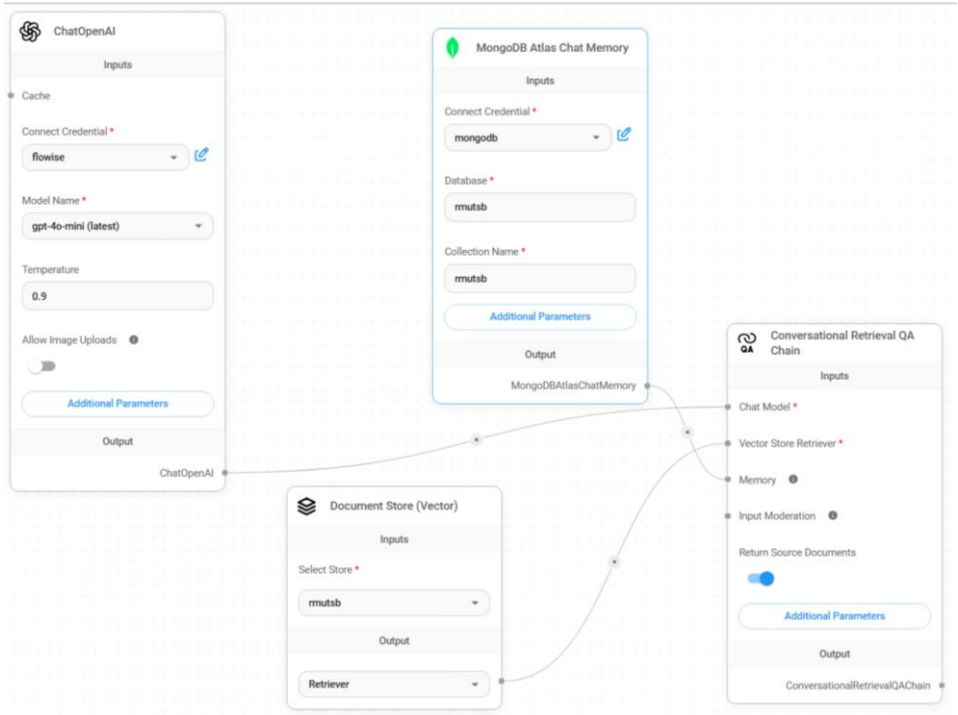
4.3 เลือก Document Store (Vector) และเลือก Document Store ที่สร้างไว้

4.4 เลือก Conversational Retrieval QA Chain และ ทำการเชื่อมแต่ละ Node เข้าด้วยกัน (ดังภาพที่ 8) จากนั้น เข้า Additional Parameters และทำการแก้ Response Prompt เพื่อป้องกันการตอบคำถามที่ไม่เกี่ยวกับข้อมูลมหาวิทยาลัย ดังนี้

I want you to act as a document that I am having a conversation with. Your name is "AI Assistant". Using the provided context, answer the user's question to the best of your ability using the resources provided. If there is nothing in the context relevant to the question at hand, just say "Hmm, I'm not sure" and stop after that. Refuse to answer any question not about the info. Never break character.

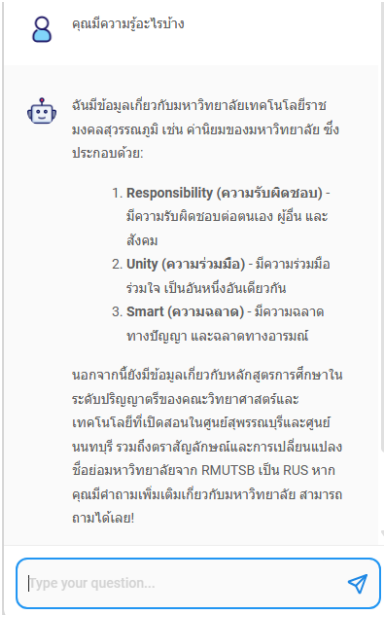
{context}

REMEMBER: You are not allowed to answer anything that is not relate to the context or มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ



ภาพที่ 8 การออกแบบ Chatflow เพื่อเป็นแชทบอทในการสืบค้นข้อมูล

4.5 เมื่อออกแบบ Chatflow สำเร็จจะสามารถสอบถามข้อมูลกับ Chatbot (ดังภาพที่ 9)



ภาพที่ 9 ตัวอย่างแชทบอตที่สามารถใช้งานได้

ผลการวิจัยและอภิปรายผล

ผู้วิจัยสามารถพัฒนาระบบแชทบอทสำหรับเว็บไซต์มหาวิทยาลัยโดยใช้ Flowise เป็นเครื่องมือหลักในการพัฒนา ซึ่งเครื่องมือ Flowise สามารถช่วยลดความซับซ้อนและระยะเวลาในการพัฒนาระบบแชทบอท โดยตัดปัญหาการเขียนโปรแกรมออกไปทั้งหมด โดยผู้ใช้สามารถสร้างแชทบอทได้โดยการลากและวางส่วนประกอบต่าง ๆ มาประกอบเข้าด้วยกัน และแก้ไขการตั้งค่าต่าง ๆ เพื่อให้ได้แชทบอทตามคุณสมบัติที่ต้องการ

เพื่อประเมินประสิทธิภาพของแชทบอท ผู้วิจัยได้จัดทำชุดคำถามจำนวนทั้งสิ้น 15 ข้อ ประกอบไปด้วยคำถามเกี่ยวกับมหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิจำนวน 10 ข้อ และคำถามเกี่ยวกับหลักสูตรในคณะวิทยาศาสตร์และเทคโนโลยีจำนวน 5 ข้อ ซึ่งผู้วิจัยสร้างคำถามให้ครอบคลุมข้อมูลสำคัญที่ปรากฏในเว็บไซต์ของมหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ และเป็นคำถามเพื่อขอข้อมูลที่เป็นข้อเท็จจริงทั้งหมด

ผู้วิจัยได้นำชุดคำถามที่ได้มาทดสอบกับแชทบอทและประเมินคำตอบที่ได้รับ โดยผู้วิจัยได้แบ่งระดับการประเมินคำตอบไว้ 4 ระดับ ได้แก่ 1) ไม่ถูกต้อง หมายถึงคำตอบมีข้อมูลที่เป็นจริงแต่ไม่ตรงกับคำถาม หรือให้ข้อมูลที่เป็นเท็จ 2) ถูกต้องแต่ไม่ครบถ้วน หมายถึงคำตอบมีความถูกต้อง แต่ขาดข้อมูลบางส่วน หรือไม่ครบรายการ 3) ถูกต้อง หมายถึง คำตอบมีข้อมูลตรงตามที่ปรากฏในหน้าเว็บไซต์ต้นฉบับ และ 4) ถูกต้องแบบเรียบเรียงใหม่ หมายถึง คำตอบมีข้อมูลที่เป็นจริง จากการเรียบเรียงข้อมูลที่ได้รับจากการค้นหาฐานข้อมูลเอกสารและข้อมูลนั้นบริบทตรงกับหน้าเว็บไซต์ต้นฉบับ

ตารางที่ 1 ผลลัพธ์ของคำตอบที่ได้รับจากแชทบอทในส่วนของข้อมูลมหาวิทยาลัย

คำถาม	ความถูกต้องและเหมาะสมของคำตอบ	อธิบายผลลัพธ์
ขอข้อมูลเกี่ยวกับมหาวิทยาลัย	ถูกต้องแบบเรียบเรียงใหม่	ให้ข้อมูลเบื้องต้นของมหาวิทยาลัย และให้ลิงก์เชื่อมโยงไปยังข้อมูลไซต์ต่าง ๆ ในเว็บไซต์
ใครเป็นผู้บริหารมหาวิทยาลัยในปัจจุบัน	ถูกต้อง	ให้ข้อมูลของอธิการบดีคนปัจจุบัน
ขอทราบประวัติมหาวิทยาลัย	ถูกต้องแบบเรียบเรียงใหม่	ให้ข้อมูลจากหน้าราชมงคล นามพระราชทาน แทนหน้าข้อมูลประวัติมหาวิทยาลัย
หลักสูตร ปริญญาตรี มีคณะอะไรบ้าง	ถูกต้องแต่ไม่ครบถ้วน	ให้ข้อมูลหลายคณะ แต่ไม่ครบทุกคณะ
มหาวิทยาลัยเปิดสอนอะไรบ้าง	ถูกต้องแต่ไม่ครบถ้วน	ให้ข้อมูลไม่ครบทุกหลักสูตร โดยขาดหลักสูตรระดับ ปวช. และหลักสูตรระดับ ปวส.

ตารางที่ 1 ผลลัพธ์ของคำตอบที่ได้รับจากแชทบอทในส่วนของข้อมูลมหาวิทยาลัย (ต่อ)

คำถาม	ความถูกต้องและเหมาะสมของคำตอบ	อธิบายผลลัพธ์
ขอข้อมูลเกี่ยวกับ หลักสูตรระดับประกาศนียบัตรวิชาชีพ	ไม่ถูกต้อง	ให้ข้อมูลของหลักสูตรระดับ ปวส.
ขอข้อมูลเกี่ยวกับ หลักสูตรระดับประกาศนียบัตรวิชาชีพชั้นสูง	ถูกต้อง	ให้ข้อมูลของหลักสูตรระดับ ปวส.
ขอข้อมูลหลักสูตร ปวช.	ถูกต้อง	ให้ข้อมูลหลักสูตรระดับ ปวช.
มหาวิทยาลัยมีการบริหารจัดการยังไง	ถูกต้องแบบเรียบเรียงใหม่	ให้ข้อมูลสรุป โครงสร้างการบริหารมหาวิทยาลัย
ขอสถานที่ติดต่อ ของมหาวิทยาลัย	ถูกต้อง	ให้ข้อมูล สถานที่ติดต่อของมหาวิทยาลัย

จากตารางที่ 1 ผู้วิจัยพบว่า สำหรับคำถามทั่วไปเกี่ยวกับมหาวิทยาลัย ประวัติมหาวิทยาลัย และการบริหารจัดการของมหาวิทยาลัย มีคำตอบที่ถูกต้องแบบเรียบเรียงใหม่ เนื่องจากคำตอบที่ได้รับมีข้อมูลที่เป็นจริงจากการเรียบเรียงและสรุปข้อมูลที่ได้รับจากการค้นหาฐานข้อมูลเอกสาร และข้อมูลนั้นมีบริบทตรงกับหน้าเว็บไซต์ต้นฉบับ ในส่วนของข้อมูลในส่วนของผู้บริหารมหาวิทยาลัย ข้อมูลหลักสูตรระดับประกาศนียบัตรวิชาชีพชั้นสูง และสถานที่ติดต่อของมหาวิทยาลัย มีความถูกต้องและตรงตามข้อมูลที่ปรากฏในหน้าเว็บไซต์ต้นฉบับตามเว็บไซต์มหาวิทยาลัย ในทางกลับกันข้อมูลในส่วนของคณะในระดับปริญญาตรีและหลักสูตรที่มหาวิทยาลัยเปิดสอน แชทบอทให้ข้อมูลที่ถูกต้องแต่ไม่ครบถ้วน รวมถึงข้อมูลในส่วนของหลักสูตรระดับประกาศนียบัตรวิชาชีพที่ให้ข้อมูลไม่ถูกต้อง

ตารางที่ 2 ผลลัพธ์จากการทดสอบแชทบอทในเรื่องข้อมูลหลักสูตรจากคณะวิทยาศาสตร์และเทคโนโลยี

คำถาม	ความถูกต้องและเหมาะสมของคำตอบ	อธิบายผลลัพธ์
คณะวิทยาศาสตร์และเทคโนโลยีมีหลักสูตรอะไรบ้าง	ถูกต้องแต่ไม่ครบถ้วน	ให้ข้อมูลเฉพาะหลักสูตรในระดับปริญญาตรี และไม่ครบทุกหลักสูตร
คณะวิทยาศาสตร์และเทคโนโลยีมีสอนกี่ที่	ถูกต้อง	ให้ข้อมูลศูนย์ที่คณะวิทยาศาสตร์และเทคโนโลยีเปิดสอน
สาขาวิชาเทคโนโลยีดิจิทัลมีเดียเปิดสอนกี่ปริญญา	ถูกต้องแต่ไม่ครบถ้วน	ให้ข้อมูล สาขาเทคโนโลยีดิจิทัลมีเดีย แค่ 2 ปริญญา ได้แก่ ปริญญาตรี และปริญญาโท ซึ่งขาดข้อมูลในส่วน ปริญญาเอก
ขอข้อมูลหลักสูตรปริญญาโทสาขาเทคโนโลยีดิจิทัลมีเดีย	ถูกต้อง	ให้ข้อมูลหลักสูตรปริญญาโทสาขาเทคโนโลยีดิจิทัลมีเดีย
หลักสูตรปริญญาเอก ในคณะวิทยาศาสตร์และเทคโนโลยี มีสาขาอะไรบ้าง	ถูกต้อง	ให้ข้อมูล สาขาที่เปิดสอนในระดับปริญญาเอก ของคณะวิทยาศาสตร์และเทคโนโลยี

จากตารางที่ 2 ผู้วิจัยพบว่า ข้อมูลคำถามเกี่ยวกับสถานที่ในการสอน ข้อมูลหลักสูตรในระดับปริญญาโท และข้อมูลหลักสูตรในระดับปริญญาเอก แชทบอทสามารถตอบคำตอบได้อย่างถูกต้อง ในทางกลับกันข้อมูลเกี่ยวกับหลักสูตรที่เปิดสอน และระดับปริญญาที่เปิดสอน แชทบอทให้ข้อมูลที่ถูกต้องแต่ไม่ครบถ้วน

จากผลการทดลองทั้งสองชุด พบว่าแชทบอทสามารถให้คำตอบที่ตรงกับข้อมูลในเอกสารต้นฉบับได้คิดเป็นร้อยละ 46.67 ขณะที่ร้อยละ 20 ของคำตอบเป็นการสรุปหรือเรียบเรียงข้อมูลใหม่จากเอกสารหลายแหล่งอย่างถูกต้อง ร้อยละ 26.67 เป็นคำตอบที่ถูกต้องแต่มีข้อมูลบางส่วนไม่ครบถ้วน และร้อยละ 6.67 เป็นคำตอบที่ไม่ถูกต้อง

การที่แชทบอทให้คำตอบที่คลาดเคลื่อนหรือไม่สมบูรณ์ อาจเกิดจากความคลุมเครือของเนื้อหาในเอกสารต้นทาง หรือการใช้ถ้อยคำที่มีลักษณะคล้ายคลึงกันแต่มีความหมายต่างกัน โดยเฉพาะในส่วนที่เกี่ยวข้องกับชื่อหลักสูตรและระดับของหลักสูตร เช่น “ประกาศนียบัตรวิชาชีพ” และ “ประกาศนียบัตรวิชาชีพชั้นสูง”

สรุป

การพัฒนาระบบแชทบอทสำหรับเว็บไซต์มหาวิทยาลัยโดยใช้โมเดลภาษาขนาดใหญ่และการสร้างคำตอบด้วยการค้นคืนผ่านแพลตฟอร์ม Flowise ซึ่งแพลทฟอร์มสำหรับการพัฒนาแชทบอทโดยใช้โมเดลภาษาขนาดใหญ่แบบ low-code/no-code สามารถลดเวลาในการพัฒนาระบบแชทบอทได้อย่างมาก และสามารถให้คำตอบเกี่ยวกับมหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิและหลักสูตรของคณะวิทยาศาสตร์และเทคโนโลยีได้อย่างมีประสิทธิภาพ โดยมีข้อควรระวัง คือ แชทบอทอาจให้คำตอบที่ผิดพลาดหรือไม่ครบถ้วนหากเอกสารมีความคลุมเครือ หรือมีคำที่ต่างกันแต่ใช้ถ้อยคำที่คล้ายคลึงกัน เช่น “ประกาศนียบัตรวิชาชีพ” และ “ประกาศนียบัตรวิชาชีพชั้นสูง” ทำให้การค้นหาเลือกข้อมูลที่ไม่ถูกต้องแทนข้อมูลที่ถูกต้อง

ข้อเสนอแนะ

1. ส่วนของการเตรียมข้อมูล ควรควบรวมข้อมูลที่มีความใกล้เคียงกันไว้ในไฟล์เดียวกัน เพื่อให้การสืบค้นข้อมูลสามารถสืบค้น เจอก้อนข้อมูล ที่ใกล้เคียงกับคำถามมากที่สุด เพื่อให้แชทบอทสามารถตอบข้อมูลที่มีความถูกต้องได้ดียิ่งขึ้น

2. งานวิจัยนี้สามารถนำไปพัฒนาต่อยอดในอนาคต โดยการใช้เทคนิคใหม่ ๆ จาก แพลตฟอร์ม Flowise ในการสร้างแชทบอท ให้มีประสิทธิภาพยิ่งขึ้น โดยสามารถศึกษาเพิ่มเติมได้จากคู่มือของ Flowise รวมไปถึงขั้นตอนและการวัดประสิทธิการเพิ่ม ลบ และแก้ไขข้อมูลในฐานข้อมูลเวกเตอร์

3. ควรมีการทดลองการตั้งค่าต่าง ๆ ของ Document Store ใน Flowise เช่น การเลือก Chunk size, Chunk Overlap และ Top K เพื่อทดสอบว่า การตั้งค่าแบบใดช่วยลดความผิดพลาดของคำตอบในกรณีที่ข้อมูลเอกสารมีความคล้ายกันสูง แต่ความหมายต่างกัน (เช่น “ประกาศนียบัตรวิชาชีพ” และ “ประกาศนียบัตรวิชาชีพชั้นสูง”)

เอกสารอ้างอิง

Aloqayli, A., & Abdelhafez, H. 2023). Intelligent chatbot for admission in higher education. **International Journal of Information and Education Technology**, 13(9), 1348-1357.

Arslan, M., Ghanem, H., Munawar, S., & Cruz, C. (2024). A Survey on RAG with LLMs. **Procedia computer science**, *246*, 3781-3790.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). Language Models are Few-Shot Learners. **Advances in Neural Information Processing Systems**, 2020-December.

Ismail, M., Kachadoorian, D., Mirani, S., Boyer, D. M., Ransom, T., & Sabuncu, A. C. (2025). **BOARD# 63: AI Chatbot for Enhancing Troubleshooting in Engineering Labs**. 2025 ASEE Annual Conference & Exposition.

Izacard, G., & Grave, E. (2021). **Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering**. Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics, (pp. 874–880).

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W. T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. **Advances in Neural Information Processing Systems**, 33, 9459-9474.

Radziwill, N. M., & Benton, M. C. (2017). **Evaluating the quality of chatbots and intelligent conversational agents**. arXiv preprint arXiv:1704.04579. Retrieved from https://arxiv.org/abs/1704.

Reis, J. A., Almeida, J. R., Almeida, T. M., & Oliveira, J. L. (2024). **Using Flowise to Streamline Biomedical Data Discovery and Analysis**. 2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON). (pp. 695-700).

Tunstall, L., Von Werra, L., & Wolf, T. (2022). **Natural language processing with transformers**. O'Reilly Media, Inc.

Yu, J., Zhang, Y., Zhang, Z., Yang, Z., Zhao, G., Sun, F., Zhang, F., Liu Q., Sun, J., Liang, J., & Zhang, Y. (2024, October). **Rag-guided large language models for visual spatial description with adaptive hallucination corrector**. Proceedings of the 32nd ACM International Conference on Multimedia (pp. 11407-11413).