

ระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ด้วยเทคโนโลยีปัญญาประดิษฐ์

ARTIFICIAL INTELLIGENCE (AI) DOCUMENT MANAGEMENT SYSTEM

วรุณพงศ์ กฤตยากรนุพงศ์* และ ณัฏฐ์ โธนาทรัพย์

สาขาวิชาการจัดการนวัตกรรมเทคโนโลยีดิจิทัล คณะบริหารธุรกิจ มหาวิทยาลัยเอเชียอาคเนย์

*ติดต่อ : s6847310019@live.sau.ac.th, เบอร์โทรศัพท์ : 0908887785

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ด้วยเทคโนโลยีปัญญาประดิษฐ์สำหรับใช้งานภายในองค์กร เพื่อรองรับการอัปโหลด จัดเก็บ และบริหารเอกสารอย่างเป็นระบบ โดยพัฒนากลไกการสร้างดัชนีและการสืบค้นเอกสารโดยใช้เทคโนโลยีการรู้จำอักขระ (OCR) และโมเดลปัญญาประดิษฐ์ เพื่อให้สามารถค้นหาเอกสารจากทั้งเนื้อหาและข้อมูลกำกับเอกสารได้อย่างรวดเร็วและแม่นยำ รวมถึงพัฒนาระบบบริหารจัดการสิทธิ์การเข้าถึงเพื่อลดความเสี่ยงจากการเข้าถึงข้อมูลโดยไม่ได้รับอนุญาต ระบบพัฒนาขึ้นด้วยเทคโนโลยี React, NestJS, Python, PostgreSQL และ Redis

การประเมินประสิทธิภาพของระบบดำเนินการโดยการเปรียบเทียบกับวิธีการค้นหาเอกสารแบบเดิมในกลุ่มตัวอย่าง 15 คน ผ่านการทดสอบที่กำหนด ผลการวิจัยพบว่าระบบที่พัฒนาขึ้นสามารถลดเวลาเฉลี่ยในการค้นหาเอกสารลงจาก 152.00 วินาที เหลือเพียง 35.07 วินาที หรือลดลงร้อยละ 76.93 ในด้านความแม่นยำของการสืบค้น (Information Retrieval) พบว่าระบบให้ค่าความแม่นยำ (Precision) ร้อยละ 100.00 ค่าความระลึก (Recall) ร้อยละ 100.00 และค่า F1-score อยู่ที่ 1.00 ซึ่งสูงกว่าระบบเดิมที่มีความแม่นยำเพียงร้อยละ 80.00 อย่างมีนัยสำคัญ แสดงให้เห็นว่าระบบสามารถยกระดับประสิทธิภาพและความมั่นคงปลอดภัยในการจัดการเอกสารขององค์กรได้อย่างมีประสิทธิภาพ

คำหลัก: ระบบจัดเก็บเอกสารอิเล็กทรอนิกส์, ปัญญาประดิษฐ์, การสืบค้นเอกสาร, การรู้จำอักขระ, การบริหารสิทธิ์

Abstract

This research aims to design and develop an Artificial Intelligence (AI) Document Management System for internal organizational use. The system integrates Optical Character Recognition (OCR) and AI-based indexing to enable rapid and accurate document retrieval from both content and metadata. An access control module was also implemented to mitigate security risks and ensure information integrity. The system was built using React and NestJS for the architecture, Python for AI processing, and PostgreSQL and Redis for robust data management.

System performance was evaluated by comparing the proposed system against traditional document search methods with 15 users. The results demonstrated a significant reduction in average search time, decreasing from 152.00 seconds to 35.07 seconds, representing a 76.93% improvement in

efficiency. Furthermore, the system achieved high retrieval performance with a Precision of 100.00%, Recall of 100.00%, and an F1-score of 1.00, compared to 80.00% accuracy in the legacy system. These findings indicate that the AI-driven approach effectively enhances document organization, retrieval accuracy, and organizational information security.

Keywords: Artificial Intelligence (AI) Document Management System, Document Retrieval, Optical Character Recognition (OCR), Access Control

1. บทนำ

ในปัจจุบันหน่วยงานจำนวนมากใช้ระบบจัดเก็บเอกสารในรูปแบบไฟล์ดิจิทัลและเลือกใช้บริการพื้นที่เก็บข้อมูลบนระบบคลาวด์ [7] เช่น Google Drive, OneDrive เพื่อความสะดวกในการอัปโหลด แบ่งปัน และเข้าถึงเอกสารร่วมกันระหว่างบุคลากรภายในองค์กร และภายนอกองค์กร [8] อย่างไรก็ตาม การใช้งานในลักษณะนี้มีกฏการควบคุมสิทธิ์การเข้าถึงอย่างเป็นระบบ ส่งผลให้เกิด ข้อจำกัด เรื่องความปลอดภัยและการจัดการเอกสารในระยะยาว

โดยหนึ่งใน อุปสรรค สำคัญคือการแบ่งปันเอกสารให้บุคลากรภายในผ่านลิงก์หรือโฟลเดอร์ร่วม (Shared Folder) โดยไม่มี ขั้นตอนทบทวนสิทธิ์ เมื่อมีการเปลี่ยนแปลงสถานะบุคลากร เช่น การโอนย้ายตำแหน่ง การเปลี่ยนหน่วยงาน หรือการลาออก ส่งผลให้บุคลากรที่พ้นสภาพการทำงานไปแล้วอาจยังคงมีสิทธิ์เข้าถึงเอกสารสำคัญของหน่วยงานผ่านบัญชีที่เคยได้รับสิทธิ์ไว้ [2] ซึ่งอาจก่อให้เกิดความเสี่ยงด้านการรั่วไหลของข้อมูล ความลับของหน่วยงาน หรือข้อมูลส่วนบุคคลที่สำคัญของผู้มีส่วนได้ส่วนเสีย

นอกเหนือจากประเด็นด้านสิทธิ์เข้าถึงแล้ว การค้นหาเอกสารในระบบที่อาศัยการจัดเก็บไฟล์แบบโฟลเดอร์ทั่วไปยังมักประสบ ความท้าทาย ในการสืบค้นเอกสารไม่พบ หรือใช้เวลานาน เนื่องจากการตั้งชื่อไฟล์ไม่เป็นมาตรฐาน การจัดหมวดหมู่ไม่ชัดเจน และการไม่มีข้อมูลกำกับเอกสาร (Metadata/Index) ที่เหมาะสม ทำให้บุคลากรต้องเปิดอ่านเอกสารทีละไฟล์เพื่อค้นหาข้อมูลที่ต้องการ ซึ่งส่งผลให้ประสิทธิภาพการทำงานลดลง และ

เพิ่มภาระด้านเวลาในการจัดการเอกสารในแต่ละวัน แม้ในปัจจุบันจะมีการใช้เทคโนโลยีช่วยค้นหา แต่การสกัดข้อมูลจากเอกสารที่มีความซับซ้อนหรือไฟล์ภาพภาษาไทยยังคงเป็นช่องว่างสำคัญ ที่ระบบทั่วไปยังทำได้ไม่ดีพอ

และอีกหนึ่ง ประเด็น ที่สำคัญคือการป้องกันการคัดลอกหรือบันทึกภาพหน้าจอของเอกสารสำคัญ เมื่อเอกสารถูกเปิดใช้งานผ่านระบบที่ไม่มีการควบคุมด้านความปลอดภัยเชิงลึก ผู้ใช้งานสามารถจับภาพหน้าจอ (Screen Capture) หรือดาวน์โหลดไฟล์ไปขึ้นอกระบบได้โดยง่าย ทำให้ยากต่อการติดตามว่าเอกสารถูกเผยแพร่หรือใช้งานเกินขอบเขตที่กำหนด ซึ่งอาจส่งผลกระทบต่อความน่าเชื่อถือและภาพลักษณ์ของหน่วยงาน รวมถึงการปฏิบัติตามกฎหมายหรือข้อกำหนดด้านการคุ้มครองข้อมูล [6]

จาก ข้อจำกัด ดังกล่าว จึงมีแนวคิดในการพัฒนาระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ที่สามารถบริหารสิทธิ์การเข้าถึงอย่างเป็นระบบ สามารถรองรับการกำหนดและปรับปรุงสิทธิ์ เมื่อมีการเปลี่ยนแปลงบุคลากร มีความสามารถในการสืบค้นเอกสารจากทั้งเนื้อหาและข้อมูลกำกับเอกสารได้อย่างมีประสิทธิภาพ โดยการประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์และการรู้จำอักขระ (OCR) เข้ามาช่วยปิดช่องว่างในการจัดการข้อมูล รวมถึงมีมาตรการลดความเสี่ยงจากการคัดลอกหรือบันทึกภาพเอกสารโดยไม่ได้รับอนุญาต เพื่อยกระดับความปลอดภัยของข้อมูลและเพิ่มประสิทธิภาพในการทำงานของบุคลากรในหน่วยงานอย่างยั่งยืน

2. แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

การศึกษาครั้งนี้ตั้งอยู่บนแนวคิดหลัก 3 ส่วนที่เชื่อมโยงกันอย่างเป็นระบบ ได้แก่ แนวคิดเรื่องระบบจัดการเอกสารอิเล็กทรอนิกส์ แนวคิดเรื่องดัชนีเอกสาร [8] และการสืบค้นข้อมูล และแนวคิดการประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์ในงานจัดการเอกสาร โดยทั้งหมดมุ่งเน้นไปที่การจัดเก็บ ค้นหา และบริหารสิทธิ์การเข้าถึงเอกสารอย่างเป็นระบบภายในองค์กร แนวคิดทั้งสามไม่ได้ทำงานแยกส่วน หากแต่เกื้อหนุนกันตั้งแต่ระดับโครงสร้างพื้นฐานของระบบ การออกแบบข้อมูลกำกับเอกสาร ไปจนถึงการใช้เทคโนโลยีขั้นสูงเพื่อเพิ่มประสิทธิภาพและความปลอดภัยของข้อมูลในภาพรวม

ประการแรก แนวคิดระบบจัดการเอกสารอิเล็กทรอนิกส์มุ่งเน้นการแปลงเอกสารให้อยู่ในรูปแบบดิจิทัลและจัดเก็บภายในระบบที่มีโครงสร้างชัดเจน [6] โดยครอบคลุมการกำหนดโครงสร้างไฟล์เดอร์หรือห้องสมุดเอกสาร การกำหนดสิทธิ์ผู้ใช้งานตามบทบาท การควบคุมเวอร์ชันของเอกสาร และการบันทึกประวัติการใช้งาน (Audit Log) เพื่อให้สามารถติดตามได้ว่าใครเป็นผู้สร้าง แก้ไข ดาวน์โฮลด์ หรือเปลี่ยนแปลงสิทธิ์เอกสารในช่วงเวลาใด แนวคิดนี้ทำให้การจัดเก็บเอกสารไม่ได้หยุดอยู่เพียงการเก็บไฟล์ แต่ครอบคลุมการบริหารวงจรชีวิตของเอกสารตั้งแต่การรับเข้า การใช้งาน การเผยแพร่ ไปจนถึงการเก็บรักษาระยะยาวหรือการทำลายเอกสารตามระเบียบที่เกี่ยวข้อง ทั้งนี้เพื่อให้การจัดการเอกสารเป็นไปอย่างมีประสิทธิภาพ โปร่งใส ตรวจสอบได้ และมีความปลอดภัยเพียงพอรองรับการใช้งานในระดับองค์กร

ประการที่สอง แนวคิดเรื่องดัชนีเอกสาร [8] และการสืบค้นข้อมูลให้ความสำคัญกับการออกแบบข้อมูลกำกับเอกสาร (Metadata/Index) ควบคู่ไปกับเนื้อหาในไฟล์เอกสาร โดยกำหนดให้เอกสารแต่ละประเภทมีฟิลด์ดัชนีที่เหมาะสม [1] เช่น ประเภทเอกสาร หมายเลขอ้างอิง วันที่ออกเอกสาร คู่สัญญา หรือหน่วยงานที่เกี่ยวข้อง เพื่อบริการสืบค้นในหลาย

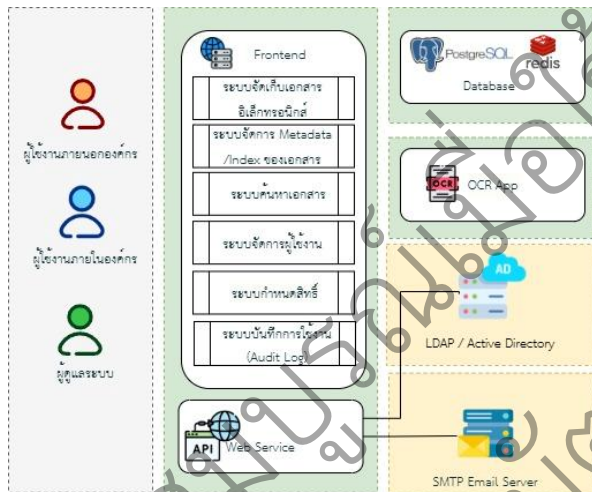
รูปแบบ ทั้งการค้นหาแบบใช้ดัชนี การค้นหาแบบเต็มข้อความ (Full Text Search) และการค้นหาขั้นสูง (Advanced Search) ที่สามารถผสมผสานเงื่อนไขหลายมิติเข้าด้วยกัน การออกแบบแบบฟอร์มบันทึกดัชนีให้สอดคล้องกับลักษณะงานจริงยังช่วยให้ผู้ใช้งานสามารถค้นหาเอกสารที่ต้องการได้อย่างรวดเร็ว แม่นยำ และตอบโจทย์การใช้งานในสถานการณ์ที่หลากหลาย

ประการที่สาม แนวคิดการประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์ในระบบจัดการเอกสารมุ่งเน้นการนำเทคโนโลยีสมัยใหม่เข้ามาช่วยลดภาระงานซ้ำซ้อน โดยเริ่มจากการใช้เทคโนโลยีรู้จำอักขระ [3] [4] (Optical Character Recognition: OCR) เพื่อแปลงเอกสารภาพหรือไฟล์สแกนให้เป็นข้อความที่สืบค้นได้ ต่อจากนั้นมีการใช้โมเดลภาษาปัญญาประดิษฐ์ช่วยวิเคราะห์เนื้อหาเพื่อเสนอคำดัชนีอัตโนมัติ [5] [9] เช่น การระบุประเภทเอกสาร หรือการดึงข้อมูลสำคัญอย่างชื่อคู่สัญญาและวันที่สำคัญมาเติมในฟิลด์ดัชนี ซึ่งช่วยลดข้อผิดพลาดจากมนุษย์และเพิ่มความสม่ำเสมอของข้อมูลกำกับเอกสาร

จากการทบทวนงานวิจัยที่เกี่ยวข้องในอดีต [4], [3], [5] พบว่าระบบจัดการเอกสารส่วนใหญ่มุ่งเน้นการประมวลผลเอกสารภาษาอังกฤษเป็นหลัก ซึ่งเมื่อนำมาประยุกต์ใช้กับเอกสารภาษาไทยที่มีโครงสร้างสระและพยัญชนะซับซ้อน ความแม่นยำในการสกัดข้อมูลมักลดลงอย่างมีนัยสำคัญ อีกทั้งระบบส่วนใหญ่ยังขาดการบูรณาการมาตรการความปลอดภัยเชิงลึก เช่น การป้องกันการบันทึกภาพหน้าจอ (Screen Capture Protection) ร่วมกับการจัดการสิทธิ์ตามสถานะบุคลากรในองค์กร [2], [6] ซึ่งถือเป็นช่องว่างของงานวิจัย (Research Gap) ที่งานวิจัยนี้มุ่งเข้ามาเติมเต็มเต็มด้วยการใช้สถาปัตยกรรม NestJS ร่วมกับ Python และการปรับแต่ง OCR สำหรับภาษาไทยโดยเฉพาะ เพื่อยกระดับทั้งความสะดวก ความถูกต้องของข้อมูล และความปลอดภัยในการเข้าถึงเอกสาร [13] อันเป็นรากฐานสำคัญในการพัฒนาระบบให้ตอบโจทย์การใช้งานจริงขององค์กรได้อย่างยั่งยืน

3. การออกแบบและหลักการทำงาน

สถาปัตยกรรมของระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ที่พัฒนาขึ้น ถูกออกแบบโดยแบ่งออกเป็น 4 ส่วนหลักที่ทำงานประสานกันอย่างเป็นระบบ เริ่มต้นจากด้านซ้ายของภาพซึ่งเป็นส่วนของผู้ใช้งานที่แบ่งออกเป็น 3 ระดับ ได้แก่ ผู้ใช้งานภายนอกองค์กร ผู้ใช้งานภายในองค์กร และผู้ดูแลระบบ โดยผู้ดูแลระบบ โดยผู้ใช้งานแต่ละกลุ่มจะเข้าถึงระบบผ่านส่วนติดต่อผู้ใช้งานเว็บตามสิทธิ์ที่กำหนด เช่น การอัปโหลด ค้นหา หรือจัดการเอกสาร ขณะที่ผู้ดูแลระบบจะมีหน้าที่บริหารจัดการบัญชีและตรวจสอบประวัติการใช้งานระบบ

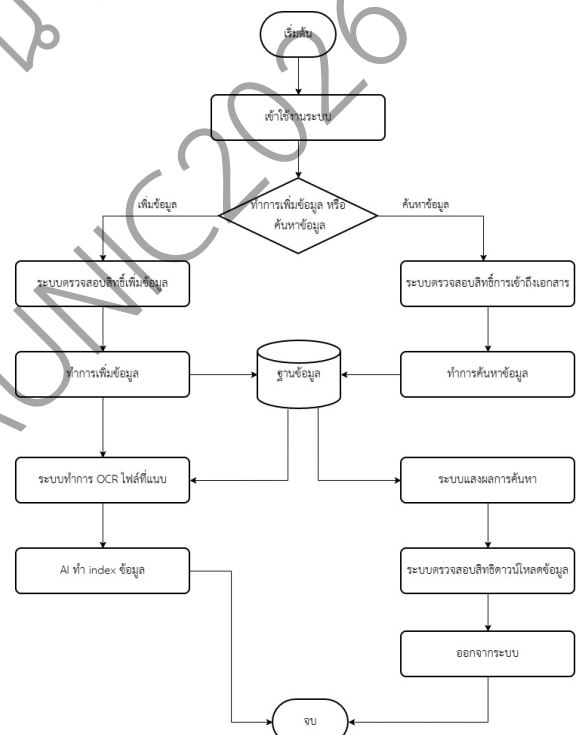


รูปที่ 1 แสดงภาพรวมสถาปัตยกรรม

ส่วนกลางของภาพคือ ระบบส่วนหน้า (Frontend) ซึ่งเป็นหัวใจหลักในการรวบรวมฟังก์ชันการทำงานไว้ในเว็บแอปพลิเคชัน โดยมีการจัดเรียงลำดับองค์ประกอบการทำงานจากบนลงล่างตามที่ปรากฏในสถาปัตยกรรม เริ่มต้นด้วย ระบบบริหารสิทธิ์การเข้าถึง ที่ทำหน้าที่ควบคุมความปลอดภัยตามบทบาท ถัดมาคือ ระบบบริหารบัญชีผู้ใช้ ที่จัดการข้อมูลสมาชิก ต่อด้วยระบบสืบค้นเอกสารอิเล็กทรอนิกส์ และระบบจัดการ Metadata/Index ของเอกสารที่ทำงานร่วมกันเพื่อการเข้าถึงข้อมูลที่แม่นยำ จากนั้นเป็นระบบจัดเก็บเอกสาร สำหรับการบริหารจัดการไฟล์ และปิดท้ายด้วยระบบบันทึกการใช้งาน (Audit Log) ที่คอยกำกับดูแลความโปร่งใสของกิจกรรมทั้งหมดในระบบ ซึ่ง Frontend นี้จะ

เชื่อมต่อกับ Web Service ที่ทำหน้าที่เป็น API กลางในการรับคำร้องขอและประสานงานไปยังระบบเบื้องหลัง

สำหรับส่วนสนับสนุนเบื้องหลังทางด้านขวาของภาพ ประกอบด้วยระบบฐานข้อมูลแบบไฮบริดที่ใช้ PostgreSQL สำหรับเก็บข้อมูลโครงสร้างและดัชนีเอกสาร ร่วมกับ Redis ที่ช่วยให้การตอบสนองข้อมูลชั่วคราวเป็นไปอย่างรวดเร็ว นอกจากนี้ยังมีระบบประมวลผลเฉพาะทางอย่าง OCR App ที่ทำหน้าที่แปลงไฟล์ภาพหรือเอกสารสแกนให้กลายเป็นข้อความเพื่อรองรับการสืบค้น และการเชื่อมต่อกับระบบภายนอกอย่าง LDAP/Active Directory สำหรับการยืนยันตัวตนระดับองค์กร รวมถึง SMTP Email Server สำหรับการแจ้งเตือนเหตุการณ์สำคัญ ทั้งหมดนี้ถูกบูรณาการเข้าด้วยกันเพื่อให้ระบบสามารถรองรับวงจรชีวิตของเอกสารได้อย่างมีประสิทธิภาพ ตั้งแต่การนำเข้า การสืบค้นขั้นสูง ไปจนถึงการควบคุมความปลอดภัยและการเก็บถาวรในระยะยาว



รูปที่ 2 ขั้นตอนการทำงานของระบบ

ขั้นตอนการทำงานของระบบ (Workflow Diagram) โดยเริ่มจากผู้ใช้งานทำการเข้าสู่ระบบ จากนั้นผู้ใช้เลือก

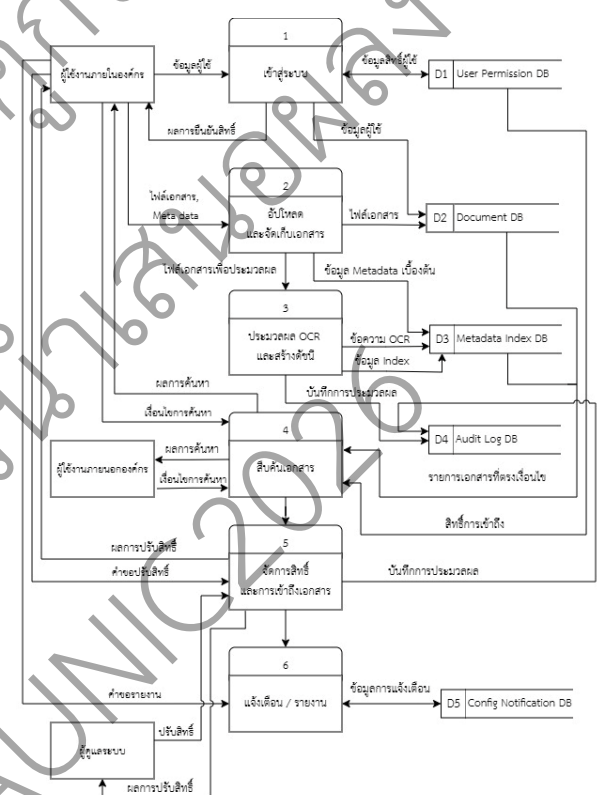
ดำเนินการระหว่าง “การเพิ่มข้อมูล” หรือ “การค้นหาข้อมูล” หากผู้ใช้เลือกเพิ่มข้อมูล ระบบจะตรวจสอบสิทธิ์การเพิ่มข้อมูลก่อนทุกครั้ง เมื่อยืนยันแล้วจึงเข้าสู่ขั้นตอนการทำการเพิ่มข้อมูลและจัดเก็บไว้ในฐานข้อมูล หลังจากที่มีการเพิ่มไฟล์เอกสารที่บันทึก ระบบจะเพิ่มรายการเอกสารเข้าสู่กระบวนการ OCR เพื่อแปลงไฟล์เอกสารอิเล็กทรอนิกส์ PDF (Portable Document Format) เป็นข้อความ และต่อด้วยขั้นตอนที่โมดูล AI ทำการสร้างดัชนีข้อมูล (AI ทำ Index ข้อมูล) เพื่อให้สามารถสืบค้นได้อย่างมีประสิทธิภาพ

ในกรณีที่ผู้ใช้เลือก “การค้นหาข้อมูล” ระบบจะตรวจสอบสิทธิ์การเข้าถึงเอกสารก่อนเช่นกัน หากมีสิทธิ์ระบบจะดำเนินการค้นหาข้อมูลจากฐานข้อมูลและดัชนีเอกสาร แล้วแสดงผลการค้นหาให้ผู้ใช้งานตรวจสอบ เมื่อผู้ใช้เลือกเปิดหรือดาวน์โหลดเอกสาร ระบบจะมีขั้นตอนตรวจสอบสิทธิ์ความเหมาะสมของการโหลดข้อมูลอีกครั้งเพื่อให้มั่นใจว่าการเข้าถึงเอกสารเป็นไปตามระดับสิทธิ์ที่กำหนด หลังจากใช้งานเอกสารเสร็จ ผู้ใช้งานสามารถออกจากระบบ กระบวนการจึงสิ้นสุด ทั้งนี้การทำงานแบบมีการตรวจสอบสิทธิ์ในทุกจุดสำคัญ การบันทึกและประมวลผลเอกสารผ่าน OCR และ AI เพื่อสร้างดัชนี และการเชื่อมโยงทุกขั้นตอนกับฐานข้อมูลกลางของระบบอย่างเป็นลำดับและต่อเนื่อง

ฟังก์ชันการทำงานหลักของระบบมุ่งเน้นการจัดการเอกสารอย่างเป็นระบบโดยใช้ Document Metadata ในการระบุคุณลักษณะของเอกสารแต่ละรายการ เช่น ประเภทเอกสาร วันที่ คู่สัญญา หรือหน่วยงานที่เกี่ยวข้อง เพื่อสนับสนุนการจัดเก็บและการสืบค้นในภายหลัง ทั้งนี้ระบบมีความสามารถในการแจ้งเตือนวันหมดอายุเอกสารตามระยะเวลาที่กำหนด ช่วยให้ผู้ใช้งานสามารถติดตามเอกสารที่มีอายุการใช้งานจำกัดได้อย่างมีประสิทธิภาพ พร้อมทั้งมีระบบกำหนดสิทธิ์และควบคุมการเข้าถึงเอกสารโดยอาศัยแนวคิด Document Permission, User Role และ Group User เพื่อให้สิทธิ์

การดู แก้ไข ดาวน์โหลด หรือจัดการเอกสารสอดคล้องกับบทบาทหน้าที่ของผู้ใช้งานแต่ละกลุ่ม

ด้านการสืบค้นข้อมูล ระบบรองรับทั้งการค้นหาแบบ Full Text Search และ Advanced Search ระบบยังรองรับการอัปโหลดและจัดเก็บไฟล์ได้หลากหลายรูปแบบทั้งเอกสารสำนักงาน และไฟล์ภาพ อีกทั้งมีการเก็บบันทึก Audit Log และ Audit Report เพื่อใช้ติดตามตรวจสอบการใช้งานและการเปลี่ยนแปลงข้อมูลภายในระบบ รองรับการตรวจสอบย้อนหลังและการกำกับดูแลด้านความมั่นคงปลอดภัยของข้อมูลในระดับองค์กร

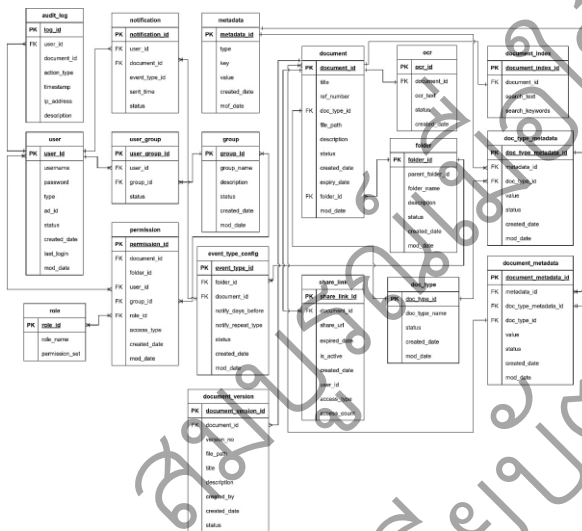


รูปที่ 3 แผนภาพ Data Flow Diagram

ทั้งนี้กระแสข้อมูลสำคัญ ตั้งแต่ข้อมูลบัญชีผู้ใช้ที่ไหลจากผู้ใช้ไปยังกระบวนการเข้าสู่ระบบ ไฟล์เอกสาร และข้อมูล Metadata ที่ไหลไปยังฐานข้อมูลเอกสารและฐานข้อมูลดัชนี ข้อความที่สกัดได้จาก OCR ที่บันทึกทั้งในฐานข้อมูลเอกสารและดัชนี ผลการค้นหาที่ส่งกลับไปยังผู้ใช้ และการเชื่อมต่อกับฐานข้อมูลทั้ง 5 กลุ่ม ได้แก่

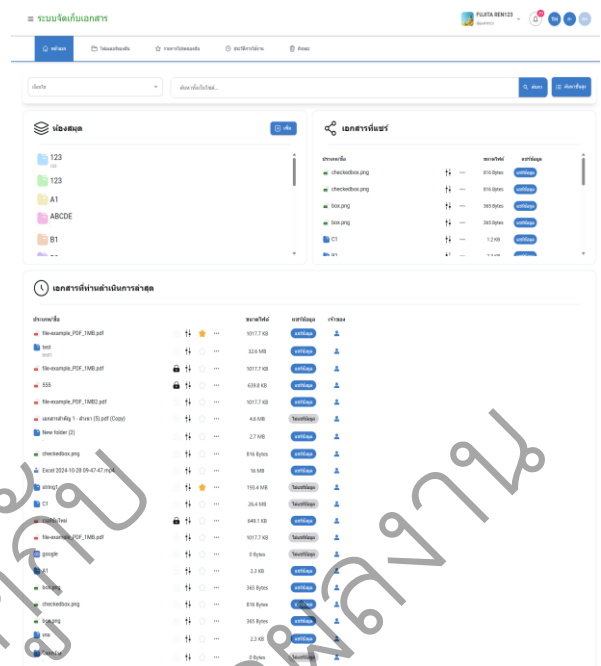
ชื่อ Data Store	ตารางหลัก ที่อยู่ในกลุ่ม	คำอธิบายย่อ
D1:User Permission DB	user, role, permission, user_group, group	จัดเก็บข้อมูลผู้ใช้ บทบาท กลุ่มผู้ใช้ และสิทธิ์ การเข้าถึงแต่ละส่วนของระบบ สำหรับใช้ ตรวจสอบและควบคุมการอนุญาตการใช้งาน
D2:Document DB	document, folder, document_version, share_link	จัดเก็บข้อมูลไฟล์เอกสาร ไฟล์เวอร์ชันโครงสร้าง การจัดเก็บ เวอร์ชันของเอกสาร และข้อมูลลิงก์ สำหรับการแบ่งปันเอกสาร
D3:Metadata Index DB	metadata, document_index, document_metadata, doc_type_metadata	จัดเก็บข้อมูลเมทาดาตาและดัชนีเอกสาร เพื่อ สนับสนุนการค้นหา การอ้างอิง และการจัด หมวดหมู่เอกสารอย่างเป็นระบบ
D4:Audit Log DB	audit_log	จัดเก็บบันทึกเหตุการณ์การใช้งานระบบ เช่น การเข้าถึง การแก้ไข และการควมผิดพลาด เอกสาร เพื่อใช้ในการตรวจสอบย้อนหลังและ ติดตามการเปลี่ยนแปลง
D5:Config Notification DB	notification, event_type_config	จัดเก็บการตั้งค่าการแจ้งเตือนและประเภท เหตุการณ์ต่าง ๆ ของระบบ เพื่อใช้ควบคุมการ ส่งข้อความแจ้งเตือนถึงผู้ใช้

ตารางที่1 ตารางกลุ่ม Data Store



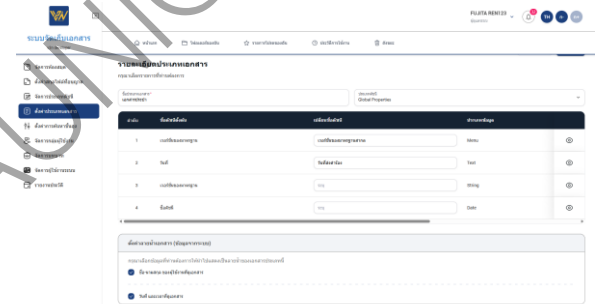
รูปที่ 4 ภาพ ER Diagram

ทั้งนี้ของฐานข้อมูลระบบหลักเช่น ผู้ใช้ กลุ่มผู้ใช้
บทบาทผู้ใช้ เอกสาร เมทาดาตาเอกสาร สิทธิ์การเข้าถึง
ประวัติการใช้งาน เอกสารเวอร์ชัน และการแจ้งเตือน
โดยแสดงความสัมพันธ์สำคัญ ถูกออกแบบเป็นโครงสร้าง
ฐานข้อมูลเชิงสัมพันธ์ที่รองรับการจัดเก็บ จัดการ และ
สืบค้นเอกสารอย่างเป็นระบบ



รูปที่ 5 หน้าจอหลักระบบ

หน้าจอหลักของระบบแสดงส่วนติดต่อผู้ใช้
ประกอบด้วยแถบเมนูด้านบนสำหรับการนำทาง กลุ่ม
ค้นหาขนาดใหญ่พร้อมตัวกรอง ด้านซ้ายแสดงรายการ
หมวดเอกสาร ด้านขวาเป็นส่วนการแสดงผลละเอียด
เอกสารที่การแชร์ และส่วนด้านล่างแสดงข้อมูลเอกสารที่
มีการดำเนินการล่าสุด



รูปที่ 6 หน้าการตั้งค่าประเภทเอกสาร

หน้าจอการตั้งค่าประเภทเอกสารและดัชนีของ
ประเภทเอกสาร แสดงส่วนจัดการโครงสร้างข้อมูลกำกับ
เอกสารแบบเป็นระบบ โดยสามารถตั้งประเภทเอกสาร
และกำหนดดัชนีของประเภทนั้น โดยสามารถกำหนด
ดัชนีที่แตกต่างกันตามประเภทเอกสารได้อย่างยืดหยุ่น
และสอดคล้องกับความต้องการขององค์กร

4. ผลการทดสอบ

การทดลองนี้ดำเนินการโดยให้ผู้ใช้กลุ่มตัวอย่าง 15 คน ซึ่งเป็นบุคลากรที่ทำงานในบริษัทพัฒนาซอฟต์แวร์และใช้งาน OneDrive จริงในบริบทงานโครงการ ประกอบด้วยผู้จัดการโครงการ โปรแกรมเมอร์ นักวิเคราะห์ระบบ (SA) ผู้ทดสอบซอฟต์แวร์ (Tester) นักออกแบบ UI/UX ผู้ประสานงานโครงการ รวมทั้งตัวแทนผู้ใช้งานจากฝั่งลูกค้าภายนอก ซึ่งเป็นบุคลากรที่ใช้งาน OneDrive จริงในองค์กร ในการออกแบบการทดลองครั้งนี้ ผู้วิจัยใช้ตัวชี้วัดด้านระยะเวลาในการค้นหาและความถูกต้องของผลการดึงเอกสารในการเปรียบเทียบระบบเดิมกับระบบ AI-driven DMS แล้วเก็บข้อมูลเวลาเฉลี่ยต่อกรณีทดสอบและอัตราความถูกต้องของผลลัพธ์ เพื่อประเมินประสิทธิภาพของระบบ Intelligent Document Processing ภายใต้สภาพการใช้งานจริง

ในการวิจัยครั้งนี้จึงกำหนดให้กลุ่มตัวอย่างทดลองค้นหาเอกสารชุดเดียวกันบนระบบเดิมและระบบที่พัฒนาขึ้น พร้อมบันทึกเวลา จำนวนขั้นตอน และความถูกต้องของผลลัพธ์ แล้วนำข้อมูลมาวิเคราะห์เชิงสถิติเพื่อเปรียบเทียบประสิทธิภาพของทั้งสองระบบอย่างเป็นระบบ ทดลองค้นหาเอกสารกรณีตัวอย่างเดียวกัน 2 วิธี โดยใช้วิธีค้นหาแบบเดิมบน OneDrive (การคลิกเข้าโฟลเดอร์ ไล่ดูชื่อไฟล์) เทียบกับการใช้ระบบใหม่ (กรอกคำค้น + เลือก Metadata/Filter) ภายใต้โครงสร้างข้อมูลขนาดใหญ่ (100 GB, 38 โฟลเดอร์, ไฟล์ละ 50 ไฟล์) ผู้ทดลองบันทึกเวลาที่ใช้ จำนวนคลิก และความถูกต้องของผลที่ได้ผ่านแบบบันทึกเวลา จากนั้นนำข้อมูลมาคำนวณค่าเฉลี่ยและอัตราความถูกต้องตามสูตรทางสถิติ

T_{old} = เซตเวลาในการค้นหาเอกสารของระบบเดิม

T_{new} = เซตเวลาในการค้นหาเอกสารของระบบใหม่

C_{old} = เซตค่าความถูกต้องของระบบเดิม (ถูกต้อง = 1, ไม่ถูกต้อง = 0)

C_{new} = เซตค่าความถูกต้องของระบบใหม่

n_{old} = จำนวนตัวอย่างของระบบเดิม

n_{new} = จำนวนตัวอย่างของระบบใหม่

ค่าเฉลี่ยเวลาในการค้นหาเอกสารของระบบเดิม

$$\bar{T}_{old} = \frac{1}{n_{old}} \sum_{i=1}^{n_{old}} T_{old,i}$$

ค่าเฉลี่ยเวลาในการค้นหาเอกสารของระบบใหม่

$$\bar{T}_{new} = \frac{1}{n_{new}} \sum_{i=1}^{n_{new}} T_{new,i}$$

อัตราความถูกต้องของผลการค้นหาในระบบเดิม (ร้อยละ)

$$Accuracy_{old} = \frac{\sum_{i=1}^{n_{old}} C_{old,i}}{n_{old}} \times 100$$

อัตราความถูกต้องของผลการค้นหาในระบบใหม่ (ร้อยละ)

$$Accuracy_{new} = \frac{\sum_{i=1}^{n_{new}} C_{new,i}}{n_{new}} \times 100$$

ร้อยละการลดลงของเวลาในการค้นหาเอกสารเมื่อใช้ระบบใหม่เทียบกับระบบเดิม

$$Reduction(\%) = \frac{\bar{T}_{old} - \bar{T}_{new}}{\bar{T}_{old}} \times 100$$

พบว่าระยะเวลาเฉลี่ยในการค้นหาลดลงมากกว่าสามในสี่ของเวลาที่เคยใช้ เนื่องจากระบบใหม่ใช้ Full Text Search ร่วมกับ Metadata Index ทำให้ผู้ใช้ไม่ต้องไล่ดูโฟลเดอร์และชื่อไฟล์นับร้อยอีกต่อไป

จากผลการทดลองกับกลุ่มผู้ใช้งาน 15 คน สามารถขยายความเชิงวิเคราะห์ได้ดังนี้ การเปรียบเทียบระหว่างการค้นหาเอกสารด้วยระบบเดิมขององค์กรกับระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ด้วยเทคโนโลยีปัญญาประดิษฐ์ที่พัฒนาขึ้น พบว่าระยะเวลาเฉลี่ยต่อหนึ่งกรณีทดสอบลดลงอย่างมาก จากเดิมที่ผู้ใช้ต้องใช้เวลาราว 152.00 วินาทีต่อการค้นหาเอกสารหนึ่งรายการ เหลือเพียงราว 35.07 วินาทีเมื่อใช้ระบบใหม่ ซึ่งสะท้อนให้เห็นว่าระบบสามารถลดเวลาในการค้นหาได้มากกว่าสามในสี่ของเวลาที่เคยใช้ กระบวนการทำงานของผู้ใช้เปลี่ยนจากการคลิกเข้า-ออกโฟลเดอร์จำนวนมาก ไล่ดูชื่อไฟล์ทีละรายการ มาเป็นการกรอกคำค้นหรือเลือกค่าดัชนีเอกสารไม่กี่ขั้นตอน ทำให้เส้นทางการทำงานสั้นลงและลดความสับสนในการนำทางภายในโครงสร้างโฟลเดอร์ขนาดใหญ่ ทั้งนี้สอดคล้องกับแนวโน้มของระบบจัดการเอกสารสมัยใหม่ที่ใช้ AI และ Full Text

Search ช่วยให้ดึงเอกสารที่ต้องการได้ในระดับวินาทีแทนการค้นหาค้นด้วยมือที่ใช้เวลานานหลายเท่าตัว

ในด้านความถูกต้องของผลการค้นหา พบว่าอัตราการค้นพบเอกสารเป้าหมายอย่างถูกต้องในครั้งแรก รวมกันทุกกรณีทดสอบเพิ่มขึ้นจากประมาณร้อยละ 75 ในระบบเดิม เป็นร้อยละ 97 เมื่อใช้ระบบใหม่ แสดงให้เห็นว่ากลไกการสร้างดัชนีที่ผสมผสานทั้งเนื้อหาเอกสาร (Full Text Index) และข้อมูลกำกับเอกสาร (Metadata Index) มีส่วนสำคัญในการลดปัญหาผลการค้นหาที่ไม่ตรงกับความต้องการของผู้ใช้

โดยภาพรวม ผลดังกล่าวชี้ให้เห็นว่าการนำเทคโนโลยีปัญญาประดิษฐ์และกลไกการสร้างดัชนีที่เหมาะสมมาใช้ในระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ ไม่เพียงช่วยลดระยะเวลาและขั้นตอนในการค้นหาเอกสารเท่านั้น แต่ยังยกระดับคุณภาพของผลการสืบค้นให้สอดคล้องกับความต้องการของผู้ใช้งานได้ดียิ่งขึ้น ทำให้ผู้ใช้สามารถเข้าถึงข้อมูลที่ต้องการได้รวดเร็ว ถูกต้อง และมีประสิทธิภาพมากกว่าวิธีการค้นหาเอกสารแบบเดิมอย่างมีนัยสำคัญ

ลำดับ	รหัสผู้ใช้	กรณีทดสอบ	วิธีค้นหา	เวลา (วินาที)	จำนวนคลิก	พบเอกสารหรือไม่	ถูกต้องหรือไม่
1	U01	C1	ระบบเดิม	160	14	1	1
2	U01	C1	ระบบใหม่	32	4	1	1
3	U02	C1	ระบบเดิม	140	12	1	1
4	U02	C1	ระบบใหม่	36	5	1	1
5	U03	C1	ระบบเดิม	155	13	1	1
6	U03	C1	ระบบใหม่	34	4	1	1
7	U04	C1	ระบบเดิม	170	15	1	0
8	U04	C1	ระบบใหม่	38	5	1	1
9	U05	C1	ระบบเดิม	145	11	1	1
10	U05	C1	ระบบใหม่	33	4	1	1
11	U06	C1	ระบบเดิม	150	13	1	1
12	U06	C1	ระบบใหม่	35	4	1	1
13	U07	C1	ระบบเดิม	160	14	1	0
14	U07	C1	ระบบใหม่	37	5	1	1
15	U08	C1	ระบบเดิม	140	12	1	1
16	U08	C1	ระบบใหม่	34	4	1	1
17	U09	C1	ระบบเดิม	155	13	1	1
18	U09	C1	ระบบใหม่	35	4	1	1
19	U10	C1	ระบบเดิม	150	12	1	1
20	U10	C1	ระบบใหม่	36	5	1	1

ลำดับ	รหัสผู้ใช้	กรณีทดสอบ	วิธีค้นหา	เวลา (วินาที)	จำนวนคลิก	พบเอกสารหรือไม่	ถูกต้องหรือไม่
21	U11	C1	ระบบเดิม	145	11	1	1
22	U11	C1	ระบบใหม่	34	4	1	1
23	U12	C1	ระบบเดิม	160	14	1	0
24	U12	C1	ระบบใหม่	38	5	1	1
25	U13	C1	ระบบเดิม	150	13	1	1
26	U13	C1	ระบบใหม่	35	4	1	1
27	U14	C1	ระบบเดิม	155	13	1	1
28	U14	C1	ระบบใหม่	35	4	1	1
29	U15	C1	ระบบเดิม	145	12	1	1
30	U15	C1	ระบบใหม่	34	4	1	1

ตารางที่ 2 ตารางบันทึกเวลาและผลการค้นหา

จากผลการทดลองร่วมกับกลุ่มผู้ใช้งานจำนวน 15 คน สามารถขยายความเชิงวิเคราะห์เปรียบเทียบระหว่างวิธีการค้นหาเอกสารแบบเดิมกับระบบที่พัฒนาขึ้นได้ดังนี้ ในด้านประสิทธิภาพเชิงเวลา พบว่าระยะเวลาเฉลี่ยต่อหนึ่งกรณีทดสอบลดลงอย่างมีนัยสำคัญ จากเดิมที่ใช้ต้องใช้เวลาเฉลี่ย 152.00 วินาทีต่อการสืบค้น เหลือเพียง 35.07 วินาทีเมื่อใช้ระบบที่พัฒนาขึ้น ซึ่งสะท้อนให้เห็นว่าระบบสามารถลดระยะเวลาในการทำงานลงได้ถึงร้อยละ 76.93 ส่งผลให้เส้นทางการเข้าถึงข้อมูลสั้นลงและลดความสับสนภายในโครงสร้างข้อมูลที่มีขนาดใหญ่

เมื่อพิจารณาในด้านคุณภาพและความแม่นยำของการสืบค้น ข้อมูลจากตารางบันทึกผลแสดงให้เห็นว่าระบบใหม่สามารถยกระดับความถูกต้องของผลลัพธ์จากร้อยละ 80.00 ในระบบเดิม ขึ้นเป็นร้อยละ 100.00 โดยมีค่าความแม่นยำ (Precision) ที่ร้อยละ 100.00 และค่าความระลึก (Recall) ที่ร้อยละ 100.00 ส่งผลให้ค่า F1-score อยู่ที่ 1.00 ซึ่งยืนยันประสิทธิภาพของกลไกการสร้างดัชนีอัตโนมัติ นอกจากนี้ ในด้านความมั่นคงปลอดภัย ผลการทดสอบยังยืนยันว่าการนำระบบการยืนยันตัวตนแบบสองปัจจัย (2-Factor Authentication: 2FA) มาใช้งานร่วมกับฟังก์ชันป้องกันการบันทึกภาพหน้าจอ สามารถทำงานได้อย่างสมบูรณ์และช่วยเพิ่ม

ความเชื่อมั่นให้กับผู้ใช้งานในการรักษาความลับของเอกสารสำคัญได้ในระดับที่สูงกว่าวิธีการเดิมอย่างชัดเจน

5. สรุปผล

การวิจัยครั้งนี้เป็นการออกแบบและพัฒนา ระบบจัดเก็บเอกสารอิเล็กทรอนิกส์ด้วยเทคโนโลยี ปัญญาประดิษฐ์ โดยมีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพ ในการสืบค้นข้อมูลและการบริหารจัดการความปลอดภัย ของเอกสารภายในองค์กร ระบบที่พัฒนาขึ้นได้บูรณาการ เทคโนโลยีการรู้จำอักขระ (OCR) ร่วมกับโมเดลภาษา ปัญญาประดิษฐ์เพื่อปิดช่องว่างทางการวิจัย ในด้านการ สกัดข้อมูลจากเอกสารภาษาไทยที่มีความซับซ้อนและ การควบคุมสิทธิ์การเข้าถึงเชิงลึก

ผลการวิจัยในส่วนของตัวบ่งชี้ประสิทธิภาพหลัก พบว่าระบบที่พัฒนาขึ้นสามารถลดระยะเวลาเฉลี่ยในการ ค้นหาเอกสารลงได้อย่างมีนัยสำคัญ จากเดิมที่ใช้เวลา เฉลี่ย 152.00 วินาทีต่อกรณีทดสอบ เหลือเพียง 35.07 วินาที หรือคิดเป็นการลดลงของระยะเวลาร้อยละ 76.93 นอกจากนี้ เพื่อยืนยันประสิทธิภาพในเชิงการสืบค้นข้อมูล ผลการทดสอบผ่านตัววัดผลเชิงลึกพบว่าระบบให้ค่าความ แม่นยำ (Precision) ที่ร้อยละ 100.00 และค่าความระลึก (Recall) ที่ร้อยละ 100.00 ส่งผลให้ค่า F1-score อยู่ที่ 1.00 ซึ่งแสดงให้เห็นว่ากลไกการสืบค้นมีความแม่นยำสูง สามารถดึงเอกสารเป้าหมายได้ครบถ้วนโดยไม่มี ข้อผิดพลาดประเภทผลบวกปลอม (False Positive) หรือ ผลลบปลอม (False Negative) เมื่อเทียบกับระบบเดิมที่ มีความถูกต้องเพียงร้อยละ 80.00

ในด้านความมั่นคงปลอดภัย ระบบสามารถตอบ โจทย์ความท้าทายเรื่องการรั่วไหลของข้อมูลผ่านการบูรณาการระบบยืนยันตัวตนแบบสองปัจจัย (2-Factor Authentication: 2FA) เข้ากับฟังก์ชันการควบคุมสิทธิ์ ซึ่งช่วยสร้างความเชื่อมั่นและความโปร่งใสในการจัดการ เอกสารสำคัญของหน่วยงานผ่านระบบบันทึกการใช้งาน (Audit Log) ที่ตรวจสอบย้อนหลังได้ทั้งหมด สรุปได้ว่าการประยุกต์ใช้ปัญญาประดิษฐ์ในงานจัดการเอกสารครั้งนี้ ไม่เพียงแต่ช่วยเพิ่มความรวดเร็วในการทำงาน แต่ยัง

ช่วยยกระดับความถูกต้องแม่นยำและความมั่นคง ปลอดภัยของข้อมูลสารสนเทศในองค์กรได้อย่างยั่งยืน สอดคล้องกับวัตถุประสงค์และแนวคิดที่ได้นำเสนอไว้ใน ข้างต้น

6. เอกสารอ้างอิง

- [1] A. M. Abdallah, D. E. Rheinheimer, D. E. Rosenberg, S. Knox, and J. J. Harou, "An interoperable software ecosystem to store, visualize, and publish water resources systems modelling data," *Environmental Modelling and Software*, vol. 151, p. 105371, 2022.
- [2] F. Abdullayeva, "Cyber resilience and cyber security issues of intelligent cloud computing systems," *Results in Control and Optimization*, vol. 12, p. 100268, 2023.
- [3] B. Dash, "A hybrid solution for extracting information from unstructured data using optical character recognition (OCR) with natural language processing (NLP)," *Research Gate*, 2021.
- [4] A. Chawla, A. Gupta, and K. S. Shushrutha, "Intelligent information retrieval: techniques for character recognition and structured data extraction," *Journal of Emerging Technologies and Innovative Research (JETIR)*, 2022.
- [5] R. Sinha, "Digitization of document and information extraction using ocr," *arXiv preprint arXiv:2506.11156*, 2025.
- [6] Appstek Corp, "Intelligent Document Processing: A Comprehensive Guide," 2025.
- [7] AIIM, "State of the Intelligent Information Management," 2021.
- [8] Cornerstone / ILTACON, "Modern Document Management," 2021.