สัมมนาวิชาการ เทคโนโลยีดิจิทัลมีเดีย ระดับบัณฑิตศึกษา ครั้งที่ 3

ระบบอัจฉริยะเพื่อให้คำแนะนำด้านประเด็นกฎหมายโรงงานโดยใช้การสร้างข้อความแบบเสริมด้วยการค้นคืน (RAG) และโมเดลภาษาขนาดใหญ่แบบหลายตัวแทน (Multi-Agent LLMs)

A Multi-Agent LLM Framework with Retrieval-Augmented Generation for Factory Legal Consultation

นายกฤษฎิ์ภวิศ สูงสว่าง รหัสนักศึกษา 167491432001
สาขาวิชาเทคโนโลยีดิจิทัลมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

บทคัดย่อ

-

คำสำคัญ

-

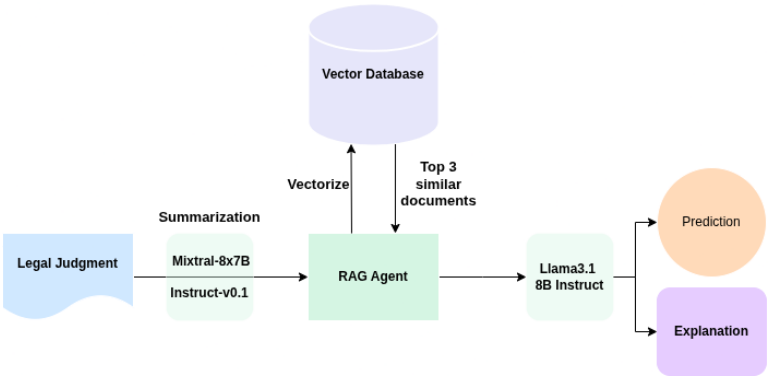
1. บทนำ

พระราชบัญญัติโรงงาน พ.ศ. 2535 [1] ถือเป็นกฎหมายแม่บทสำคัญในการกำกับดูแลภาคอุตสาหกรรมของประเทศไทย โดยมีวัตถุประสงค์เพื่อควบคุมการประกอบกิจการให้เป็นไปอย่างมีระเบียบเรียบร้อย ปลอดภัย และไม่ก่อให้เกิดผลกระทบต่อสิ่งแวดล้อม อย่างไรก็ตาม ในทางปฏิบัติกลับพบปัญหาความท้าทายในการบังคับใช้กฎหมายจากความซับซ้อนของกฎหมายและรายละเอียดทางเทคนิคจำนวนมาก ซึ่งส่งผลกระทบต่อทั้งเจ้าหน้าที่ผู้ปฏิบัติงานและผู้ประกอบการ ในมุมมองของเจ้าหน้าที่รัฐ จากการลงพื้นที่สำรวจและสัมภาษณ์เจ้าหน้าที่ผู้ปฏิบัติงานที่มีความเชี่ยวชาญเฉพาะด้าน พบว่าอุปสรรคสำคัญคือความไม่คล่องตัวในการสืบค้นข้อกฎหมายที่มีความกระจัดกระจายเพื่อนำมาวินิจฉัยหน้างานได้อย่างทันท่วงที ในขณะเดียวกัน ในมุมมองของผู้ประกอบการกิจการโรงงาน ความซับซ้อนของข้อกฎหมายมักนำไปสู่ความรู้เท่าไม่ถึงการณ์หรือความเข้าใจที่คลาดเคลื่อนในรายละเอียดของข้อบังคับ ส่งผลให้เกิดข้อบกพร่องในการปฏิบัติตามมาตรฐานโดยมิได้เจตนา ดังปรากฏในกรณีศึกษาที่สะท้อนถึงผลกระทบจากการขาดความระมัดระวังในรายละเอียดของกฎหมาย อาทิ กรณีคำสั่งปิดโรงงานแห่งหนึ่งในจังหวัดสมุทรปราการ ที่เกิดข้อผิดพลาดในกระบวนการผลิตจนส่งผลกระทบต่อชุมชน [2] กรณีโรงงานหลอมตะกั่วในจังหวัดพิษณุโลก ที่พบปัญหาการจัดการกากอุตสาหกรรมที่ไม่เป็นไปตามข้อกำหนดการกำจัดของเสียอันตราย [3] และกรณีการตรวจพบความผิดปกติของมาตรฐานการประกอบกิจการในโรงงานรีไซเคิลแห่งหนึ่งในจังหวัดฉะเชิงเทรา [4] เหตุการณ์เหล่านี้สะท้อนให้เห็นว่า การขาดเครื่องมือที่ช่วยให้เข้าถึงและเข้าใจข้อกฎหมายได้อย่างถูกต้องแม่นยำ ไม่เพียงแต่เป็นภาระของเจ้าหน้าที่ แต่ยังเป็นสาเหตุสำคัญที่ทำให้ภาคอุตสาหกรรมไม่สามารถปฏิบัติตามกฎหมายได้อย่างครบถ้วนสมบูรณ์

เพื่อเป็นการแก้ปัญหาข้อจำกัดด้านการเข้าถึงและประมวลผลข้อมูลกฎหมายที่มีความซับซ้อนดังกล่าว เทคโนโลยีปัญญาประดิษฐ์ยุคใหม่จึงเข้ามามีบทบาทสำคัญในการเป็นเครื่องมือสนับสนุน โดยเฉพาะ

อย่างยิ่ง โมเดลภาษาขนาดใหญ่ (Large Language Models: LLMs) ซึ่งเป็นเทคโนโลยีที่ถูกพัฒนาขึ้นให้สามารถเรียนรู้จากคลังข้อมูลมหาศาล เพื่อทำความเข้าใจและสังเคราะห์ภาษาของมนุษย์จนสามารถนำไปประยุกต์ใช้ได้อย่างแพร่หลาย [5] อีกทั้งศักยภาพของ LLMs ได้รับการพิสูจน์แล้วในหลากหลายสาขาที่ต้องอาศัยความแม่นยำของข้อมูล อาทิ การตอบคำถามหรือวิเคราะห์ข้อมูลทางด้านการเงิน [6] การให้คำแนะนำด้านการเกษตร [7] ไปจนถึงการสนับสนุนข้อมูลทางการแพทย์ [8] ความสำเร็จในสาขาต่างๆ เหล่านี้นำไปสู่ความสนใจของนักวิจัยในการนำ LLMs มาประยุกต์ใช้กับงานด้านกฎหมาย ซึ่งเป็นศาสตร์ที่ต้องใช้ทรัพยากรข้อมูลจำนวนมากเช่นกัน โดยผลจากการศึกษาวิจัยพบว่า โมเดลภาษาขนาดใหญ่สามารถทำหน้าที่เป็นผู้ช่วยทางด้านกฎหมาย สนับสนุนให้เจ้าหน้าที่ทำงานได้รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น แม้ว่าจะยังไม่สามารถทำหน้าที่ตัดสินใจแทนมนุษย์ได้ทั้งหมดก็ตาม [9] ด้วยเหตุนี้ จึงกล่าวได้ว่าโมเดลภาษาขนาดใหญ่เป็นเครื่องมือที่มีศักยภาพสูงและเหมาะสมอย่างยิ่ง ที่จะนำมาพัฒนาเป็นระบบสนับสนุนเจ้าหน้าที่ในการสืบค้นและให้คำปรึกษาด้านกฎหมายโรงงานแก่ผู้ประกอบการได้อย่างถูกต้องและแม่นยำ

อย่างไรก็ตาม การนำโมเดลภาษาขนาดใหญ่มาประยุกต์ใช้ในการให้คำปรึกษาทางด้านกฎหมายที่ต้องอาศัยความถูกต้องสูงนั้น จำเป็นต้องมีกระบวนการเชื่อมโยงความสามารถทางภาษาเข้ากับฐานข้อมูลอ้างอิง เพื่อลดความเสี่ยงจากการสร้างข้อมูลเท็จ และแก้ไขปัญหาข้อมูลที่ไม่เป็นปัจจุบัน โดยเทคนิคที่ได้รับการยอมรับในการแก้ปัญหาดังกล่าวคือ กระบวนการที่เรียกว่า Retrieval-Augmented Generation (RAG) [10] ทั้งนี้ การนำ กระบวนการ Retrieval-Augmented Generation ไปประยุกต์ใช้นั้นสามารถดำเนินการได้ด้วยเทคนิคที่แตกต่างกันขึ้นอยู่กับรูปแบบข้อมูลและวิธีการของผู้วิจัย โดยจากการทบทวนวรรณกรรม พบว่า การจัดการกับเอกสารกฎหมายที่มีความยาวและซับซ้อนมักประสบปัญหาข้อจำกัดด้านหน้าต่างบริบท ของโมเดล ทำให้งานวิจัยในปัจจุบันเริ่มหันมาใช้เทคนิคการลดทอนขนาดข้อมูล ดังเช่นในระบบ NyayaRAG [11] ที่ใช้วิธีการสรุปย่อเนื้อหาของคำพิพากษาและข้อกฎหมาย แทนการใช้เอกสารฉบับเต็มในการจัดนำเข้าข้อมูลสู่ฐานข้อมูลเวกเตอร์และสืบค้น เพื่อให้โมเดลสามารถรับข้อมูลในปริมาณที่เหมาะสมได้ แต่จุดอ่อนสำคัญของแนวทางที่พึ่งพาการย่อความเช่นนี้คือ การสูญเสียรายละเอียดเชิงลึก อย่างถาวรตั้งแต่ขั้นตอนการเตรียมข้อมูล ซึ่งอาจทำให้ประเด็นละเอียดอ่อนทางกฎหมาย ข้อยกเว้นเฉพาะ หรือเหตุผลที่ซับซ้อนตกหล่นไป ส่งผลให้โมเดลไม่สามารถเข้าถึงบริบทต้นฉบับเพื่อวินิจฉัยข้อเท็จจริงได้อย่างครบถ้วนสมบูรณ์



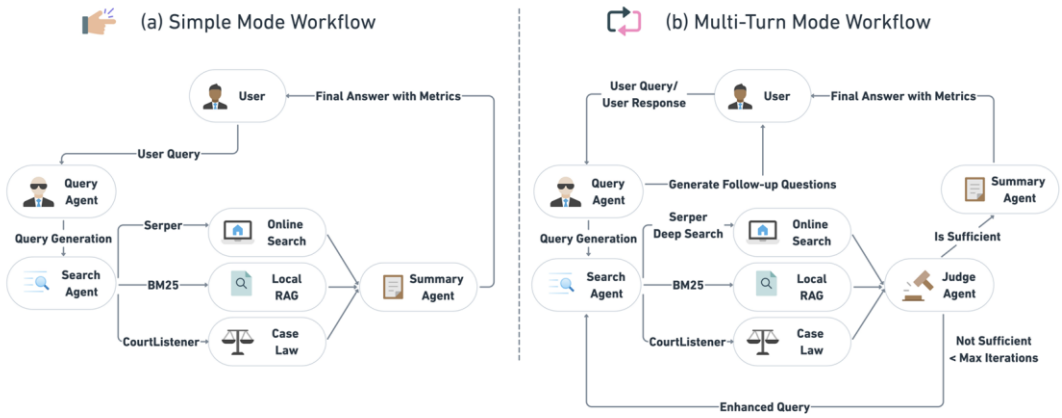
ภาพที่ 1 การทำงานของระบบ NyayaRAG [11]

สอดคล้องกับข้อค้นพบของงานวิจัย [12] ที่ได้ทำการถอดบทเรียนจากการพัฒนาระบบ RAG ในสถานการณ์จริง โดยระบุถึงจุดอ่อน 7 ข้อ ที่มักเกิดขึ้นตลอดกระบวนการพัฒนาระบบ ตั้งแต่ขั้นตอนการเตรียมนำเข้าข้อมูลสู่ฐานข้อมูลเวกเตอร์ไปจนถึงการสร้างคำตอบ ปัญหาสำคัญที่พบคือความยากในการจัดการกับบริบทของเอกสาร ซึ่งหากกำหนดขนาดการหั่นข้อมูล ไม่เหมาะสม จะส่งผลกระทบโดยตรงต่อประสิทธิภาพของระบบ กล่าวคือ หากหั่นข้อมูลเล็กเกินไปจะทำให้สูญเสียบริบทที่เชื่อมโยงกัน แต่ถ้าหั่นข้อมูลใหญ่เกินไปก็จะทำให้ได้รับข้อมูลเกินความจำเป็น ซึ่งทำให้โมเดลสับสน นอกจากนี้ระบบ RAG แบบดั้งเดิมยังขาดกลไกการตรวจสอบความถูกต้องของข้อมูลที่สืบค้นมาได้ (Retrieval Verification) ซึ่งอาจนำไปสู่การนำข้อมูลที่ไม่เกี่ยวข้องมาใช้ในการประมวลผลคำตอบ

เพื่อจัดการจุดอ่อนของระบบ RAG แบบดั้งเดิมที่ทำงาน ที่ต้องอาศัยทักษะในการเตรียมข้อมูลเพื่อนำเข้าสู่ฐานข้อมูลเวกเตอร์ รวมถึงปัญหาที่เกินจากการสืบค้นข้อมูล เพียงครั้งเดียว ไม่เพียงพอต่อการสร้างคำตอบที่แม่นยำ แนวคิดการวิจัยในปัจจุบันจึงได้มีการพัฒนาระบบที่มีสถาปัตยกรรมซับซ้อนและมีความยืดหยุ่นสูงขึ้น โดยเปลี่ยนวิธีการจากการพึ่งพาโมเดลเดียว และการสืบค้นครั้งเดียวไปสู่การใช้ ระบบหลายตัวแทน (Multi-Agent Systems) เพื่อกระจายภาระงานและตรวจสอบความถูกต้องซึ่งกันและกัน โดยมีงานวิจัยที่เกี่ยวข้อง ดังนี้:

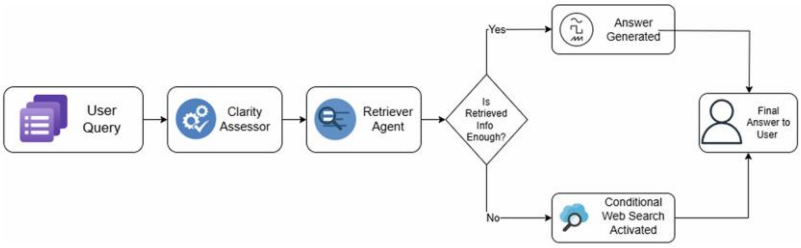
Wang และ Yuan [13] ได้เสนอระบบ L-MARS ซึ่งใช้ วิธีการให้หาคำตอบหรือสืบค้นแบบวนซ้ำ โดยระบบประกอบด้วยการทำงานประสานกันของ 3 ตัวแทน ได้แก่ (1) Query Agent ทำหน้าที่วิเคราะห์และแตกคำถามของผู้ใช้เป็นคำถามย่อย (2) Search Agent ทำหน้าที่สืบค้นข้อมูลจากแหล่งข้อมูลที่หลากหลาย และ (3) Judge Agent (ตัวแทนตัดสิน) ซึ่งทำหน้าที่ตรวจสอบคุณภาพของ

ข้อมูลที่สืบค้นมาได้ใน 3 เงื่อนไข ได้แก่ ความเพียงพอของหลักฐาน ความถูกต้องของเขตอำนาจศาล และความเป็นปัจจุบันของกฎหมาย โดยหาก Judge Agent พบว่าข้อมูลยังไม่ผ่านเกณฑ์ จะส่งสัญญาณ ให้ระบบวนกลับไปปรับปรุงคำค้นหาและสืบค้นใหม่ จนกว่าจะได้ข้อมูลที่สมบูรณ์จึงจะส่งต่อให้สร้างคำตอบสุดท้าย



ภาพที่ 2 การทำงานของระบบ L-MARS [13]

ในขณะเดียวกัน Raiyan และคณะ [14] ได้เสนอกระบวนการแก้ปัญหาความจำเพาะของกฎหมายท้องถิ่นในประเทศบังกลาเทศ โดยพัฒนาระบบที่มีสถาปัตยกรรมแบบสองตัวแทน ที่แยกหน้าที่ชัดเจน ได้แก่ (1) ตัวแทนตรวจสอบความชัดเจน ทำหน้าที่โต้ตอบกับผู้ใช้งานเพื่อ "สอบถามความชัดเจน" หากพบว่าคำถามมีความกำกวมหรือไม่ระบุรายละเอียดเพียงพอ ก่อนส่งต่อให้ (2) ตัวแทนสืบค้นข้อมูลและสร้างคำตอบ ซึ่งภายในประกอบด้วยโมเดลย่อยที่แยกหน้าที่ระหว่างการสืบค้นเอกสาร และการให้คะแนนความเกี่ยวข้อง เพื่อคัดกรองเอกสารขยะ และมีระบบค้นหาผ่านเว็บในกรณีที่ข้อมูลภายในไม่เพียงพอ ซึ่งกระบวนการนี้ช่วยลดข้อผิดพลาดจากการตีความคำถามที่คลุมเครือและเพิ่มความแม่นยำในการอ้างอิงกฎหมาย



ภาพที่ 3 การทำงานของงานวิจัยที่ [14]

ถึงแม้ว่างานวิจัยทั้งสองจะเสนอการทำงานแบบ Multi-Agent ที่มีความซับซ้อนและแม่นยำกว่าระบบเดิม แต่เมื่อจะนำเทคนิคดังกล่าวมาประยุกต์ใช้งาน กฎหมายโรงงานของไทย วิธีดังกล่าวยังมีข้อจำกัด ใน 3 เรื่องหลัก คือ

ประการแรก ขาดกลไกการจัดการความสัมพันธ์เชิงโครงสร้างของกฎหมาย: ระบบในงานวิจัย [13] และ [14] ใช้วิธีการสืบค้นข้อมูลโดยยึดตามความคล้ายคลึงของเนื้อหา เป็นหลัก ซึ่งเน้นเป็นการนำกฎหมายทุกฉบับมากองรวมกันโดยไม่แยกที่มา ส่งผลให้ระบบอาจหยิบยก กฎระเบียบย่อยหรือบทความทั่วไป มาใช้ตอบคำถามโดยละเลย กฎหมายแม่บท ที่สำคัญ ทำให้คำวินิจฉัยขาดความหนักแน่นและอาจเกิดความคลาดเคลื่อนในการอ้างอิงลำดับของกฎหมาย

ประการที่สอง ขาดความสามารถในผสมข้อมูลเชิงวิศวกรรม: กฎหมายโรงงานไม่ได้มีเพียงข้อความทางกฎหมาย แต่มีข้อมูลเชิงปริมาณและตารางมาตรฐานทางวิศวกรรม เช่น ค่าพิกัดมลพิษ, ขนาดหม้อไอน้ำ เป็นต้น ซึ่งระบบ หมายตัวแทน ทั่วไปในงานวิจัยก่อนหน้ายังไม่มี ตัวแทน เฉพาะทางที่ถูกออกแบบมาเพื่อตรวจสอบความถูกต้องของข้อมูลตารางเทคนิคเหล่านี้ร่วมกับข้อกฎหมาย

ประการที่สาม ข้อจำกัดของการทำงานแบบเรียงลำดับ: งานวิจัยก่อนหน้ามีการทำงานในลักษณะตรวจสอบตามลำดับขั้น ซึ่งจะเกิดคอขวดเมื่อต้องวินิจฉัยโรงงานหนึ่งแห่งที่เกี่ยวข้องกับกฎหมายหลายเนื้อหาพร้อมกัน เช่น ความปลอดภัย, สิ่งแวดล้อม, และเครื่องจักร ซึ่งในทางปฏิบัติจำเป็นต้องใช้การทำงานแบบขนาน เพื่อเร็วและเข้าถึงข้อมูลที่ครอบคลุม

จากข้อจำกัดของงานวิจัยก่อนหน้า ผู้วิจัยจึงมุ่งเน้นที่จะพัฒนาสถาปัตยกรรมใหม่ที่ออกแบบมาเพื่อแก้ไขข้อจำกัดเหล่านี้ โดยมีส่วนร่วมในการทำงานดังต่อไปนี้:

1. การออกแบบสถาปัตยกรรม หลายตัวแทน ที่สอดคล้องกับโครงสร้างกฎหมายไทย: นำเสนอการแยกการทำงานระหว่าง ตัวแทน กฎหมายแม่บท และ ตัวแทนกฎหมายเฉพาะทาง หรือกฎหมายลำดับรอง เพื่อให้การสืบค้นและการรวบรวมคำตอบ สามารถให้น้ำหนักความสำคัญของกฎหมายแต่ละลำดับชั้นได้อย่างถูกต้อง
2. พัฒนาวิธีการค้นคืนแบบยืดหยุ่นและปรับเปลี่ยนได้: นำเสนอแนวทางการจัดการข้อมูลที่เหมาะสมกับลักษณะเฉพาะของเอกสารกฎหมายโรงงาน โดยผสมวิธี การแบ่งส่วนข้อมูลตามมาตรา สำหรับกฎหมายแม่บท, การใช้บริบทเอกสารเต็มฉบับ สำหรับกฎหมายที่ให้ข้อมูลแบบกว้าง, การแบ่งส่วนข้อมูลระดับเอกสารทั้งฉบับ สำหรับกฎหมายเฉพาะประเภทโรงงาน

และ การแบ่งส่วนข้อมูล สำหรับตารางข้อมูลทางเทคนิค นอกจากนี้ยังเก็บเวกเตอร์แบบแยกกลุ่ม เพื่อให้ระบบสามารถประมวลผลข้อมูลที่มีความหลากหลายได้อย่างมีประสิทธิภาพสูงสุด

3. ขั้นตอนการตรวจสอบและกรองข้อมูลและ ความเกี่ยวข้องแบบพร้อมกัน: สร้างวิธีการที่ให้ตัวแทนหลายตัวทำงานตรวจสอบความเชื่อมโยง รวมถึงการทำความสะอาดผลลัพธ์ของแต่ละตัวแทน ไปพร้อมกัน และมี ตัวแทนกลางทำหน้าที่รวบรวม เพื่อให้ได้คำตอบที่ครอบคลุมทุกมิติของการประกอบกิจการโรงงาน ลดระยะเวลาการประมวลผล และเพิ่มความแม่นยำในการวินิจฉัย

จากเหตุผลและที่มาดังกล่าว ผู้วิจัยจึงเล็งเห็นถึงความสำคัญในการพัฒนาระบบอัจฉริยะเพื่อให้คำแนะนำด้านประเด็นกฎหมายโรงงาน โดยรวบรวมเทคนิคการสร้างข้อความแบบเสริมด้วยการค้นคืนแบบปรับเปลี่ยนได้ ร่วมกับสถาปัตยกรรมโมเดลภาษาขนาดใหญ่แบบหลายตัวแทน (Multi-Agent LLMs) ตามแนวทางที่นำเสนอ เพื่อสร้างเครื่องมือสนับสนุนการตัดสินใจที่มีความแม่นยำ สอดคล้องกับบริบทกฎหมายไทย และสามารถตรวจสอบแหล่งที่มาของข้อมูลได้อย่างมีประสิทธิภาพ

2. วัตถุประสงค์

2.1. เพื่อพัฒนาและออกแบบสถาปัตยกรรมระบบอัจฉริยะเพื่อให้คำแนะนำด้านประเด็นกฎหมายโรงงาน โดยบูรณาการเทคโนโลยีการสร้างข้อความแบบเสริมด้วยการค้นคืน (RAG) เข้ากับโมเดลภาษาขนาดใหญ่แบบหลายตัวแทน (Multi-Agent LLMs) ที่มีการแบ่งหน้าที่ตามความเชี่ยวชาญเฉพาะด้าน

2.2. เพื่อประเมินความถูกต้อง และความสมบูรณ์ ของคำแนะนำทางกฎหมายที่ระบบสร้างขึ้น

2.3. เพื่อศึกษาความพึงพอใจของเจ้าหน้าที่ผู้ปฏิบัติงานที่มีต่อการใช้งานระบบในการสนับสนุนการสืบค้นข้อมูลกฎหมายโรงงาน

3. ขอบเขตของการวิจัย

3.1 ประชากร

1. ประชากรด้านข้อมูลและเอกสาร: คือ กฎหมาย ระเบียบ ข้อบังคับ และเอกสารราชการทั้งหมดที่เกี่ยวข้องกับการกำกับดูแลโรงงานในประเทศไทย ภายใต้การดูแลของกรมโรงงานอุตสาหกรรมและกระทรวงอุตสาหกรรม ซึ่งครอบคลุมพระราชบัญญัติโรงงาน กฎกระทรวง ประกาศกระทรวง และหนังสือหารือต่างๆ ที่ยังมีผลบังคับใช้

2. ประชากรด้านผู้ใช้งานและผู้เชี่ยวชาญ: คือ เจ้าหน้าที่รัฐสังกัดกรมโรงงานอุตสาหกรรม, เจ้าหน้าที่สำนักงานอุตสาหกรรมจังหวัด, วิศวกรโรงงาน, และผู้ประกอบกิจการโรงงานที่มีหน้าที่เกี่ยวข้องกับการปฏิบัติตามกฎหมายโรงงาน

3.2 กลุ่มตัวอย่าง

1. กลุ่มผู้เชี่ยวชาญด้านเนื้อหา: ได้แก่ ข้าราชการระดับผู้เชี่ยวชาญหรือระดับชำนาญการพิเศษ สังกัดกรมโรงงานอุตสาหกรรม ทำหน้าที่ สร้างชุดข้อสอบและกรณีศึกษา ที่มีความซับซ้อนตามความเป็นจริง เพื่อนำมาใช้ทดสอบความแม่นยำในการประมวลผลของระบบ
2. กลุ่มผู้ประเมินความพึงพอใจ: ได้แก่ เจ้าหน้าที่ผู้ปฏิบัติงานสังกัดกรมโรงงานอุตสาหกรรม, สำนักงานอุตสาหกรรมจังหวัด หรือผู้เกี่ยวข้อง ทำหน้าที่ทดลองใช้งานระบบจริง และ ตอบแบบประเมินความพึงพอใจ ในด้านการใช้งานและประสิทธิภาพของระบบ

3.3 ตัวแปร ตัวแปรที่ใช้ได้แก่

3.3.1 ตัวแปรอิสระ

- สถาปัตยกรรมของระบบอัจฉริยะ ซึ่งประกอบด้วยการทำงานร่วมกันของ:
  1. เทคนิคการสร้างข้อความแบบเสริมด้วยการค้นคืน (Retrieval-Augmented Generation: RAG)
  2. โมเดลภาษาขนาดใหญ่แบบหลายตัวแทน (Multi-Agent LLMs) ที่มีการแบ่งหน้าที่ตามความเชี่ยวชาญ

3.3.2 ตัวแปรตาม

1. ประสิทธิภาพของระบบ ประกอบด้วย
   a. ความถูกต้อง ของคำแนะนำเมื่อเทียบกับผู้เชี่ยวชาญ
   b. ความสมบูรณ์ ของเนื้อหากฎหมายที่ครอบคลุม
2. ความพึงพอใจของผู้ใช้งาน

3.4 ระยะเวลาในการวิจัย

12 เดือน

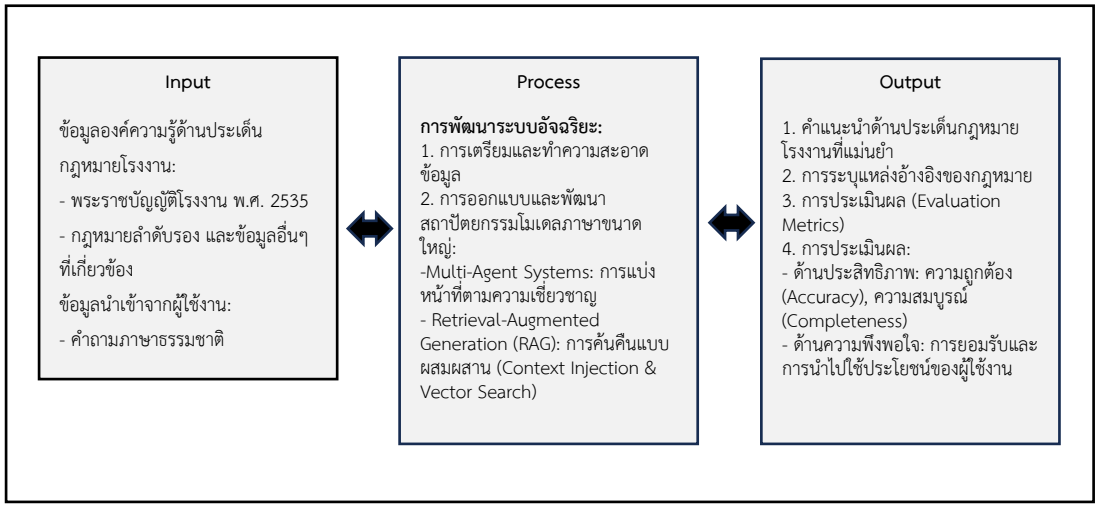
4. ประโยชน์ที่คาดว่าจะได้รับ

4.1 ได้ต้นแบบสถาปัตยกรรมระบบปัญญาประดิษฐ์ที่ผสานเทคนิคโมเดลภาษาแบบหลายตัวแทน (Multi-Agent) ร่วมกับการค้นคืนข้อมูล (RAG) ซึ่งมีประสิทธิภาพในการประมวลผลข้อมูลกฎหมายไทยที่มีความซับซ้อน และสามารถรองรับรูปแบบข้อมูลที่หลากหลาย ทั้งเอกสารข้อความ ตารางข้อมูล และรูปภาพแผนผัง

4.2 ได้เครื่องมืออัจฉริยะสำหรับสนับสนุนการปฏิบัติงานของเจ้าหน้าที่กรมโรงงานอุตสาหกรรมและหน่วยงานที่เกี่ยวข้อง ช่วยให้การสืบค้นและวินิจฉัยข้อกฎหมายเป็นไปอย่างรวดเร็ว แม่นยำ และเป็นมาตรฐานเดียวกัน ลดความผิดพลาดในการปฏิบัติหน้าที่ และช่วยยกระดับความปลอดภัยในภาคอุตสาหกรรม

4.3 เป็นแนวทางในการนำเทคโนโลยีดิจิทัลมาประยุกต์ใช้เพื่อพัฒนาระบบการกำกับดูแลกฎหมายของประเทศไทย ช่วยเพิ่มประสิทธิภาพการให้บริการภาครัฐให้มีความทันสมัย รวดเร็ว และตอบสนองความต้องการของภาคอุตสาหกรรมในยุคดิจิทัลได้อย่างมีประสิทธิภาพ

5. กรอบแนวคิดการวิจัย



ภาพที่ 5 กรอบแนวคิดของการวิจัย

จากภาพที่ 5 แสดงกรอบแนวคิดการวิจัยซึ่งจัดทำขึ้นตามรูปแบบปัจจัยนำเข้า-กระบวนการ-ผลลัพธ์ (Input-Process-Output Model) โดยมีรายละเอียดองค์ประกอบดังนี้

1. ปัจจัยนำเข้า (Input)

ประกอบด้วยแหล่งข้อมูลและวัตถุดิบสำคัญ 2 ส่วน ได้แก่

- ข้อมูลองค์ความรู้ด้านประเด็นกฎหมายโรงงาน: เป็นฐานข้อมูลหลักของระบบที่ครอบคลุมเนื้อหาพระราชบัญญัติโรงงาน พ.ศ. 2535 รวมถึงกฎหมายลำดับรองและเอกสารทางเทคนิคที่เกี่ยวข้อง เพื่อใช้เป็นคลังความรู้

- ข้อมูลนำเข้าจากผู้ใช้งาน: คือคำถามที่ระบุความต้องการในรูปแบบภาษาธรรมชาติ (Natural Language) ซึ่งสะท้อนถึงปัญหาหรือประเด็นข้อสงสัยที่เกิดขึ้นจริงในการปฏิบัติงาน

2. กระบวนการ (Process)

เป็นการดำเนินงานเพื่อสร้างนวัตกรรมตามระเบียบวิธีวิจัย โดยมุ่งเน้น การพัฒนาระบบอัจฉริยะ ซึ่งประกอบด้วยขั้นตอนสำคัญ ดังนี้

- การเตรียมและทำความสะอาดข้อมูล: กระบวนการแปลงเอกสารกฎหมายให้อยู่ในรูปแบบดิจิทัลที่พร้อมใช้งาน และจัดโครงสร้างข้อมูลให้เหมาะสม
- การออกแบบและพัฒนาสถาปัตยกรรมโมเดลภาษา: เป็นขั้นตอนการบูรณาการเทคโนโลยีปัญญาประดิษฐ์ขั้นสูง ได้แก่
    - Multi-Agent Systems: การจัดสร้างตัวแทนปัญญาประดิษฐ์หลายตัวเพื่อแบ่งหน้าที่รับผิดชอบตามความเชี่ยวชาญเฉพาะด้านของกฎหมาย
    - Retrieval-Augmented Generation (RAG): การประยุกต์ใช้เทคนิคการค้นคืนแบบผสมผสาน ทั้งการแทรกบริบท (Context Injection) สำหรับกฎหมายหลัก และการค้นหาแบบเวกเตอร์ (Vector Search) สำหรับกฎหมายเฉพาะทาง เพื่อลดข้อผิดพลาดและเพิ่มความแม่นยำ

3. ผลลัพธ์ (Output)

คือสิ่งที่ได้จากการประมวลผลของระบบและการวัดผลสัมฤทธิ์ของการวิจัย ได้แก่

- **ผลผลิตของระบบ:** คือคำแนะนำด้านประเด็นกฎหมายโรงงานที่มีความแม่นยำ พร้อมทั้งระบุแหล่งอ้างอิงของกฎหมายที่ชัดเจน เพื่อสร้างความเชื่อมั่นในการนำไปใช้งาน
- **ผลการประเมิน:** แบ่งออกเป็น 2 มิติ คือ
    - **ด้านประสิทธิภาพ:** วัดจากความถูกต้อง (Accuracy) และความสมบูรณ์ (Completeness) ของคำตอบเมื่อเทียบกับผู้เชี่ยวชาญ
    - **ด้านความพึงพอใจ:** วัดระดับการยอมรับและการนำไปใช้ประโยชน์จริงของเจ้าหน้าที่ผู้ปฏิบัติงานและกลุ่มเป้าหมาย

6. วิธีดำเนินการวิจัย

เพื่อให้การวิจัยบรรลุวัตถุประสงค์ตามที่กำหนดไว้ ผู้วิจัยได้กำหนดกรอบแนวทางการดำเนินงานโดยแบ่งออกเป็น 3 ระยะหลัก ดังนี้

6.1 ระยะที่ 1: การรวบรวมและเตรียมข้อมูลกฎหมายและข้อมูลเทคนิค

ขั้นตอนแรกมุ่งเน้นการสร้างฐานข้อมูลที่มีคุณภาพสูงและเป็นปัจจุบัน จากแหล่งข้อมูลที่มีความหลากหลายรูปแบบ เพื่อนำไปใช้เป็นฐานความรู้ให้กับโมเดลภาษา โดยมีรายละเอียดดังนี้

1. การรวบรวมข้อมูล
   รวบรวมเอกสารราชการและเอกสารทางเทคนิคทั้งหมดที่เกี่ยวข้องกับกฎหมายโรงงาน ซึ่งจัดเก็บในรูปแบบไฟล์ดิจิทัลที่หลากหลาย ได้แก่

- เอกสารข้อความ: เช่น ไฟล์ PDF ของพระราชบัญญัติโรงงาน กฎกระทรวง ประกาศกระทรวง และหนังสือหารือราชการ
- ข้อมูลตารางและฐานข้อมูล: เช่น ไฟล์ Excel (.xlsx, .xls, .csv) ที่ระบุค่ามาตรฐานทางเทคนิค บัญชีรายชื่อสารเคมีอันตราย หรือตารางพิกัดอัตราโรงงาน
- ข้อมูลบนเว็บไซต์: เช่น ไฟล์ HTML หรือหน้าเว็บเพจจากฐานข้อมูลกฎหมายออนไลน์และประกาศประชาสัมพันธ์ของหน่วยงานราชการ
- ข้อมูลรูปภาพ: เช่น ไฟล์รูปภาพ (.jpg, .png) ของแผนผังโรงงาน ตัวอย่างการติดตั้งเครื่องจักร หรือผังกระบวนการผลิต ที่แนบท้ายกฎหมาย

2. การแปลงและทำความสะอาดข้อมูล
   - การจัดการเอกสาร PDF: ดำเนินการแปลงไฟล์ ให้อยู่ในรูปแบบข้อความ พร้อมทั้งทำความสะอาดข้อมูล โดยการตัดส่วนหัวกระดาษ และท้ายกระดาษ ที่ซ้ำซ้อนในแต่ละหน้าออก เพื่อลดปริมาณข้อมูลขยะ
   - การจัดการข้อมูลตารางและเว็บ:
     - สำหรับไฟล์ Excel: ดำเนินการแปลงโครงสร้างตาราง ให้อยู่ในรูปแบบข้อความที่มีโครงสร้างชัดเจน เช่น รูปแบบ Markdown หรือ JSON เพื่อรักษาความสัมพันธ์ของข้อมูลระหว่างแถวและคอลัมน์ ให้โมเดลสามารถตีความค่ามาตรฐานต่างๆ ได้อย่างถูกต้อง
     - สำหรับไฟล์ HTML: ดำเนินการสกัดเฉพาะส่วนเนื้อหาหลัก โดยกำจัด HTML Tags, Scripts และส่วนประกอบที่ไม่เกี่ยวข้องออก
   - การจัดการข้อมูลรูปภาพ:
     - สำหรับรูปภาพที่มีตัวอักษร: ใช้เทคโนโลยี OCR (Optical Character Recognition) เพื่อแปลงข้อความในภาพให้เป็นข้อมูลดิจิทัล
     - สำหรับรูปภาพเชิงแผนภาพ: ใช้โมเดลวิเคราะห์ สร้างคำอธิบายภาพ เพื่อให้ระบบเข้าใจบริบทของรูปภาพนั้นๆ
3. การเชื่อมโยงข้อมูล
   เชื่อมโยงเนื้อหาที่ถูกตัดขาดระหว่างหน้า ให้มีความต่อเนื่อง และจัดระเบียบข้อมูลที่แปลงมาจากทุกแหล่ง เช่น PDF, Excel, HTML และรูปภาพ ให้อยู่ในรูปแบบมาตรฐานเดียวกัน เพื่อเตรียมพร้อมสำหรับกระบวนการแบ่งข้อมูล ในขั้นตอนต่อไป

6.2 ระยะที่ 2: การออกแบบและพัฒนาสถาปัตยกรรมระบบ

ดำเนินการออกแบบระบบโดยบูรณาการเทคโนโลยี Multi-Agent LLMs ร่วมกับ RAG โดยเน้นกลยุทธ์การจัดการบริบท ให้เหมาะสมกับประเภทและความสำคัญของกฎหมาย ดังนี้

1. การพัฒนาตัวแทนควบคุมและตรวจสอบกระบวนการ
   ออกแบบตัวแทนที่ทำหน้าที่เป็นสมอง ในการตัดสินใจ ตรวจสอบ และควบคุมทิศทางของคำตอบ ประกอบด้วย:

- ตัวแทนตรวจสอบขอบเขตบริบทโรงงาน: ทำหน้าที่เป็นด่านแรกในการตรวจสอบว่าคำถามของผู้ใช้มีความเกี่ยวข้องกับกฎหมายโรงงานหรือไม่ หากเป็นคำถามทั่วไปที่ไม่เกี่ยวข้อง ระบบจะปฏิเสธการตอบเพื่อลดการใช้ทรัพยากร
- ตัวแทนวิเคราะห์ความเชื่อมโยงกฎหมายลำดับรอง: ทำหน้าที่วิเคราะห์ความสัมพันธ์ว่าคำถามนั้น ต้องเชื่อมโยงไปหากฎหมายลูก (กฎกระทรวง/ประกาศกระทรวง) ฉบับใดบ้าง เพื่อส่งคำสั่งต่อไปยัง ตัวแทน เฉพาะด้านได้อย่างถูกต้อง
- ตัวแทนคัดกรองความเกี่ยวข้องของผลลัพธ์: ทำหน้าที่ตรวจสอบคำตอบเบื้องต้นที่ได้จากตัวแทน ผู้เชี่ยวชาญแต่ละตัว เพื่อตัดเนื้อหาที่ไม่เกี่ยวข้อง หรือข้อมูลที่ขัดแย้งกันออกไปก่อนส่งเข้าสู่กระบวนการรวมคำตอบ
- ตัวแทนสังเคราะห์และรวบรวมคำตอบ: ทำหน้าที่รวบรวมข้อมูลที่ผ่านการคัดกรองแล้วมาเรียบเรียงใหม่ และอ้างอิงแหล่งที่มาอย่างถูกต้องเพื่อตอบกลับผู้ใช้งาน

2. การพัฒนาตัวแทนผู้เชี่ยวชาญเฉพาะด้าน

ออกแบบตัวแทนที่ทำหน้าที่เป็น "ผู้ปฏิบัติงาน" ที่มีความรู้ลึกในแต่ละด้านและปฏิบัติงานแบบขนาน ได้แก่:

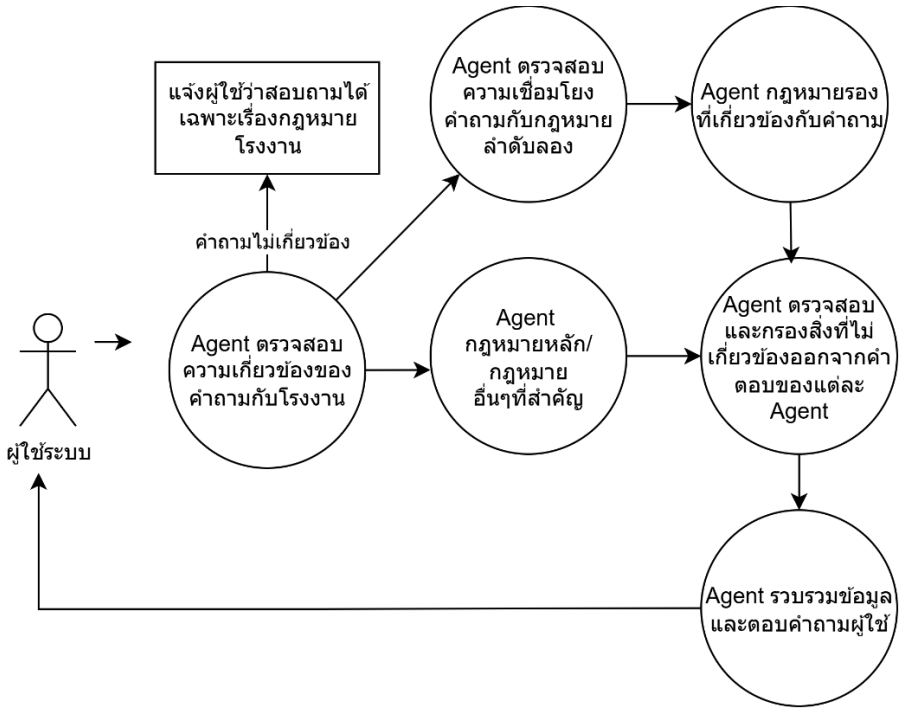
- ตัวแทนสืบค้นกฎหมายเชิงโครงสร้างและความหมาย:
  รับผิดชอบกฎหมายแม่บทที่มีความซับซ้อน (เช่น พ.ร.บ. โรงงาน) โดยตัวแทนนี้ใช้เทคนิค "การแบ่งส่วนข้อมูลตามโครงสร้างความหมาย" ซึ่งเป็นอัลกอริทึมที่ผู้วิจัยออกแบบขึ้นเฉพาะ เพื่อทำการแยก "ส่วนเนื้อหาหลัก" ออกจาก "บริบทส่วนหัวและท้าย" อย่างชาญฉลาด ทำให้ระบบสามารถสืบค้นและดึงข้อมูลรายมาตรา ออกมาประมวลผลได้โดยที่ยังคงความหมายและบริบทของกฎหมายนั้นๆ ไว้อย่างครบถ้วนสมบูรณ์
- กลุ่มตัวแทนรายด้าน: รับผิดชอบกฎหมายลำดับรอง (กฎกระทรวง/ประกาศกระทรวง/ประกาศกรม/ระเบียบหรือประกาศอื่นๆ) ที่มีขอบเขตชัดเจน เช่น ตัวแทน ด้านมลพิษทางน้ำ, ตัวแทนด้านมลพิษทางอากาศ, ตัวแทนด้านกากอุตสาหกรรม โดยตัวแทนกลุ่มนี้จะ ไม่มีการค้นหา แต่จะถูกป้อนข้อมูลกฎหมายที่เกี่ยวข้องทุกฉบับในหมวดนั้นๆ เข้าสู่หน่วยความจำโดยตรง เพื่อให้ประมวลผลความเชื่อมโยงได้ทันที
- ตัวแทนสืบค้นกฎหมายเฉพาะประเภทโรงงาน: เป็นตัวแทนหนึ่งตัวที่เชื่อมต่อกับฐานข้อมูลกฎหมายโรงงานแยกตามประเภท เช่น โรงงานน้ำตาล และ โรงงานเหล็ก เป็น

ต้น โดยตัวแทนนี้จะใช้ ระบบค้นคืน เพื่อสืบค้นว่าผู้ใช้ถามถึงโรงงานประเภทใด แล้วดึงกฎหมายฉบับที่เกี่ยวข้องมาประมวลผล

3. กลยุทธ์การทำ สืบค้น และการจัดการบริบท

ผู้วิจัยได้ออกแบบกลยุทธ์การจัดการข้อมูลออกเป็น 4 รูปแบบ ตามลักษณะของเอกสารและการใช้งาน ดังนี้:

- กลยุทธ์ RAG แบบแบ่งส่วนตามมาตรา: ใช้สำหรับ กฎหมายแม่บท (พ.ร.บ.) ที่มีขนาดใหญ่ โดยนำเทคนิคการหั่นข้อมูลตามมาตรา มาจัดทำดัชนี เพื่อให้ระบบค้นคืนเฉพาะมาตราที่เกี่ยวข้อง
- กลยุทธ์การใช้ข้อมูลฉบับเต็ม: ใช้สำหรับ ตัวแทน รายด้าน (เช่น น้ำ, อากาศ) ซึ่งมีจำนวนกฎหมายที่ใช้อ้างอิงแน่นอน ระบบจะใช้วิธีนำเข้าเอกสารกฎหมายที่เกี่ยวข้องทุกฉบับเข้าสู่คำสั่ง ของ ตัวแทน โดยตรงโดยไม่ผ่านกระบวนการค้นหาข้อมูลในเวกเตอร์ เพื่อให้ ตัวแทน มีความรู้ครอบคลุมทุกฉบับและเชื่อมโยงข้อมูลข้ามฉบับได้
- กลยุทธ์การค้นคืนระดับเอกสาร: ใช้สำหรับ ตัวแทนกฎหมายเฉพาะประเภทโรงงาน ซึ่งมีกฎหมายจำนวนมากแยกย่อยตามประเภทโรงงาน ระบบจะจัดทำดัชนีแบบ "1 ฉบับ = 1 ชิ้นข้อมูล" เมื่อผู้ใช้สอบถามเกี่ยวกับโรงงานประเภทใด ระบบจะทำการค้นหา (RAG) และดึงกฎหมายฉบับนั้นทั้งฉบับมาให้ Agent ประมวลผล
- กลยุทธ์การค้นคืนระดับแถวข้อมูล: ใช้สำหรับ ตารางข้อมูลทางเทคนิคขนาดใหญ่ (เช่น บัญชีรายชื่อสารเคมี) ระบบจะใช้เทคนิคการดึงข้อมูลเฉพาะ "แถว" ที่เกี่ยวข้องพร้อม "หัวตาราง" มาประมวลผล เพื่อลดปริมาณข้อมูลรบกวนและประหยัดทรัพยากร



ภาพที่ 4 การออกแบบสถาปัตยกรรมหลายตัวแทน (Multi Agent System)

6.3 ระยะที่ 3: การทดสอบและประเมินผลระบบ

ดำเนินการทดสอบประสิทธิภาพของระบบใน 2 มิติหลัก ได้แก่ ความถูกต้องของข้อมูล และความพึงพอใจของผู้ใช้งาน

1. การประเมินด้านประสิทธิภาพและความถูกต้อง
    - ร่วมมือกับผู้เชี่ยวชาญ จากกรมโรงงานอุตสาหกรรม ในการกำหนดชุดคำถามทดสอบและกรณีศึกษา ที่มีความซับซ้อนและครอบคลุมประเด็นสำคัญ
    - ทดสอบระบบโดยเปรียบเทียบคำตอบที่ได้กับธงคำตอบของผู้เชี่ยวชาญ เพื่อวัดค่าความถูกต้อง และความครบถ้วนสมบูรณ์ ของเนื้อหากฎหมาย
2. การประเมินความพึงพอใจของผู้ใช้งาน
    - นำระบบไปทดลองใช้งานจริงกับกลุ่มเป้าหมาย ได้แก่ เจ้าหน้าที่ผู้ปฏิบัติงานและผู้เกี่ยวข้อง
    - เก็บรวบรวมข้อมูลผ่านแบบประเมินความพึงพอใจ เพื่อศึกษาความคิดเห็นในด้านความสะดวกในการใช้งาน, ความรวดเร็วในการสืบค้นข้อมูล และประโยชน์ที่ได้รับจากการนำระบบไปใช้สนับสนุนการปฏิบัติงาน

7. ผลการวิจัย (ถ้ามี)

-

8. สรุปแนวคิดการวิจัย/ผลการวิจัย

งานวิจัยนี้มีแนวคิดหลักมุ่งเน้นการแก้ปัญหาความซับซ้อน ความกระจัดกระจาย และรายละเอียดทางเทคนิคของกฎหมายโรงงาน ซึ่งเป็นอุปสรรคสำคัญต่อการปฏิบัติงานของเจ้าหน้าที่รัฐและการปฏิบัติตามกฎหมายของผู้ประกอบการ โดยผู้วิจัยนำเสนอแนวทางการพัฒนานวัตกรรม **"ระบบอัจฉริยะเพื่อให้คำแนะนำด้านประเด็นกฎหมายโรงงาน"** ผ่านการบูรณาการเทคโนโลยีปัญญาประดิษฐ์ขั้นสูง 2 ส่วนหลัก ได้แก่ **การสร้างข้อความแบบเสริมด้วยการค้นคืน (RAG)** และ **โมเดลภาษาขนาดใหญ่แบบหลายตัวแทน (Multi-Agent LLMs)**

โดยมีสาระสำคัญของแนวคิดการวิจัย ดังนี้:

1. **การเปลี่ยนจากการใช้โมเดลเดี่ยว สู่สถาปัตยกรรมแบบหลายตัวแทน (Multi-Agent Architecture):** แทนที่จะพึ่งพาโมเดลภาษาเพียงตัวเดียว งานวิจัยนี้เสนอสถาปัตยกรรมที่จำลองกระบวนการทำงานเป็นทีมของผู้เชี่ยวชาญ โดยแบ่งตัวแทนปัญญาประดิษฐ์ (Agents) ให้รับผิดชอบเฉพาะด้าน เช่น ตัวแทนกฎหมายแม่บท, ตัวแทนกฎหมายเฉพาะทาง (น้ำ/อากาศ/กาก) และตัวแทนตรวจสอบข้อมูลเชิงวิศวกรรม เพื่อให้สามารถประมวลผลกฎหมายที่มีลำดับศักดิ์และบริบทต่างกันได้อย่างแม่นยำและทำงานคู่ขนานกันได้
2. **กลยุทธ์การค้นคืนแบบปรับเปลี่ยนได้ (Adaptive Retrieval Strategies):** เพื่อปิดจุดอ่อนของระบบ RAG แบบดั้งเดิมที่มักสูญเสียบริบทเมื่อตัดแบ่งข้อมูล งานวิจัยนี้จึงออกแบบวิธีการจัดการข้อมูลให้เหมาะสมกับประเภทเอกสาร ได้แก่ การแบ่งส่วนตามมาตรา (สำหรับ พ.ร.บ.), การใช้บริบทเอกสารเต็มฉบับ (สำหรับกฎหมายเฉพาะด้าน), การค้นคืนระดับเอกสาร (สำหรับประเภทโรงงาน) และการค้นคืนระดับแถวข้อมูล (สำหรับตารางมาตรฐานเทคนิค)
3. **การตรวจสอบความถูกต้องแบบหลายชั้น:** ระบบถูกออกแบบให้มีกลไกการตรวจสอบความเกี่ยวข้อง (Relevance Check) และความถูกต้องของข้อมูล (Verification) โดยตัวแทนกลาง ก่อนที่จะสังเคราะห์เป็นคำตอบสุดท้าย เพื่อลดปัญหาระบบสร้างข้อมูลเท็จ (Hallucination)

ผลลัพธ์ที่คาดหวังจากแนวคิดนี้ คือการได้ต้นแบบระบบสนับสนุนการตัดสินใจที่มีประสิทธิภาพสูง สามารถช่วยให้เจ้าหน้าที่กรมโรงงานอุตสาหกรรมสืบค้นและวินิจฉัยข้อกฎหมายได้อย่างรวดเร็ว ถูกต้องตามหลักวิศวกรรมและข้อกฎหมายไทย นำไปสู่การยกระดับความปลอดภัยและมาตรฐานการกำกับดูแลภาคอุตสาหกรรมของประเทศ

9. เอกสารอ้างอิง เรียงลำดับตามการกล่าวถึงในบทความ

[1] พระราชบัญญัติโรงงาน พ.ศ. 2535. (2535, 9 เมษายน). ราชกิจจานุเบกษา. เล่ม 109 ตอนที่ 44, หน้า 62-81.

[2] กรุงเทพธุรกิจ. (2566, 21 กันยายน). 'ธนกร' สั่ง ปิด 2 โรงงานผิด พ.ร.บ. ย้ำห้ามทำประชาชนเดือดร้อน. สืบค้นเมื่อ 28 มกราคม 2569,

https://www.bangkokbiznews.com/business/economic/1204114

[3] เดลินิวส์. (2567, 4 กุมภาพันธ์). สั่งฟัน! โรงงานหลฆอมตะกั่วไทย ไชน่า ซุกกากอุตสาหกรรม 3 หมื่นตัน ชี้ฝ่าฝืนต้องเร่งแก้ด่วน. สืบค้นเมื่อ 28 มกราคม 2569, https://www.dailynews.co.th/news/5541737/

[4] กรมโรงงานอุตสาหกรรม. (2567, 12 มกราคม). กรมโรงงานฯ ลุยตรวจโรงงานรีไซเคิลฉะเชิงเทรา สั่งเร่งแก้ไขด่วน! หลังพบความผิดปกติ. สืบค้นเมื่อ 28 มกราคม 2569, https://www.diw.go.th/webdiw/pr68-635/

[5] Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mimi, M. M. J., Ahmad, J., Ali, M. E., & Azam, S. (2024). A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. IEEE Access, 12, 26839-26874. https://doi.org/10.1109/ACCESS.2024.3365742

[6] Li, Y., Wang, S., Ding, H., & Chen, H. (2023). Large language models in finance: A survey. In Proceedings of the 4th ACM International Conference on AI in Finance (ICAIF '23) (pp. 374–382). ACM. https://doi.org/10.1145/3604237.3626869

[7] Biswas, S. S. (2023). Importance of chat GPT in agriculture: According to chat GPT [Unpublished manuscript]. Department of Radiology, Le Bonheur Hospital, The University of Tennessee Health Science Center.

[8] Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digital Health, 2(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

[9] Lai, J., Gan, W., Wu, J., Qi, Z., & Yu, P. S. (2023). Large language models in law: A survey. arXiv. https://doi.org/10.48550/arXiv.2312.03718

[10] Wu, Y., Wang, C., Gumusel, E., & Liu, X. (2024). Knowledge-infused legal wisdom: Navigating LLM consultation through the lens of diagnostics and positive-unlabeled reinforcement learning. arXiv. https://doi.org/10.48550/arXiv.2406.03600

[11] Nigam, S. K., Patnaik, B. D., Mishra, S., Thomas, A. V., Shallum, N., Ghosh, K., & Bhattacharya, A. (2025). *NyayaRAG: Realistic legal judgment prediction with RAG under the Indian common law system*. In K. Inui, S. Sakti, H. Wang, D. F. Wong, P. Bhattacharyya, B. Banerjee, A. Ekbal, T. Chakraborty, & D. P. Singh (Eds.), *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics* (pp. 1709–1726). The Asian Federation of Natural Language Processing and The Association for Computational Linguistics.

[12] Barnett, S., Kurniawan, S., Thudumu, S., Brannelly, Z., & Abdelrazek, M. (2024). Seven failure points when engineering a retrieval augmented generation system. In Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering—Software Engineering for AI (CAIN 2024) (pp. 194–199). ACM.

[12] Wang, Z., & Yuan, B. (2025). L-MARS: Legal multi-agent workflow with orchestrated reasoning and agentic search (arXiv:2509.00761)

[13] Jiang, C., & Yang, X. (2025). AgentsBench: A multi-agent LLM simulation framework for legal judgment prediction. Systems, 13, Article 641. https://doi.org/10.3390/systems13080641

[14] Al Raiyan, A., Zidan, M. S. H., Mehnaz, R., & Rahman, R. M. (2025). A multi-agent RAG system for legal information retrieval in Bangladesh. In S. Palaiahnakote, R. Palit, M. Saraee, P. K. Atrey, X. Bai, & B. Raman (Eds.), Data Science, AI and Applications (ICDSAIA 2025) (pp. 373–388). Communications in Computer and Information Science, vol. 2683. Springer, Cham.