

ระบบอัตโนมัติสำหรับการสกัดข้อมูลส่วนหัวการหยั่งธรณีหลุมเจาะ AUTO WELL LOG HEADER INFORMATION EXTRACTION

สุภาพร การย์บรรจบ¹
เอกสิทธิ์ พัทธวงษ์ศักดิ์²

บทคัดย่อ

ตามแผนการพัฒนาพื้นที่และสัญญาแบ่งปันผลผลิตปิโตรเลียมในอ่าวไทย จำนวนหลุมพัฒนาการผลิตมีแนวโน้มที่จะเพิ่มขึ้นอย่างรวดเร็วถึงประมาณ 500 หลุมต่อปี การจัดทำรายงานผลการประกอบการปิโตรเลียมตามมาตรฐานของกรมเชื้อเพลิงธรรมชาติในปัจจุบันต้องใช้เวลาอันยาวนานและการตรวจสอบที่ซับซ้อน ซึ่งเสี่ยงต่อข้อผิดพลาดจากการทำงานด้วยมือ โดยงานวิจัยนี้นำเสนอการใช้เทคโนโลยี Optical Character Recognition (OCR) และการสร้างโมเดล AI ผ่าน AI Builder เพื่อสกัดข้อมูลจากไฟล์ PDF ที่มีทั้งข้อความและรูปภาพ โดยโมเดล AI ที่พัฒนาขึ้นจะช่วยให้สามารถดึงข้อมูลส่วนหัวของการหยั่งธรณีหลุมเจาะ (Well Log Header) ออกมาเป็นข้อความที่สามารถนำมาจัดการและจัดรูปแบบในตาราง Excel ที่สามารถจัดการและตรวจสอบได้ง่าย ผลการวิจัยพบว่าโมเดล AI ที่พัฒนาขึ้นมีความถูกต้อง แม่นยำและมีประสิทธิภาพดี โมเดลเหล่านี้ได้ถูกนำไปใช้งานในระบบอัตโนมัติที่ออกแบบด้วย Power Automate โดยสามารถแยกประเภทข้อมูลและจัดรูปแบบได้อย่างมีประสิทธิภาพ เพื่อเพิ่มประสิทธิภาพในการสร้างและปรับปรุงโมเดล AI ควรมีการอัปเดตเอกสารรูปแบบใหม่และปรับปรุงโมเดลอย่างสม่ำเสมอเพื่อให้สอดคล้องกับการเปลี่ยนแปลงของรูปแบบเอกสารและข้อมูลอย่างแม่นยำ

คำสำคัญ : การสกัดข้อมูล, การจัดการข้อมูล, ระบบอัตโนมัติ

¹ นักศึกษาหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมข้อมูลขนาดใหญ่

² ที่ปรึกษาสารนิพนธ์หลัก

ABSTRACT

According to the development plan for areas and production sharing contracts for petroleum in the Gulf of Thailand, the number of production wells is expected to increase rapidly to approximately 500 wells per year. Currently, preparing petroleum performance reports according to the Department of Mineral Fuels' standards is a time-consuming and complex process, with a high risk of errors due to manual handling. This research presents the use of Optical Character Recognition (OCR) technology and the creation of AI models through AI Builder to extract data from PDF files containing both text and images. The developed AI models aim to extract well log header information into text that can be managed and formatted into Excel tables for easy handling and verification. The research findings indicate that the developed AI models have high accuracy and efficiency. These models have been implemented in an automated system designed with Power Automate, which efficiently categorizes and formats the data. This system enhances the speed and accuracy of data extraction and report verification. To further improve the development and refinement of AI models, it is essential to update new document formats and continuously refine the models to align with changes in document formats and data.

Keywords: Optical Character Recognition, AI Builder, Power Automate, Data Extraction

บทนำ

ในบริบทของการพัฒนาอุตสาหกรรมปิโตรเลียมในอ่าวไทย การขยายตัวของจำนวนหลุมพัฒนาการผลิตมีความสำคัญอย่างยิ่งต่อการเติบโตของเศรษฐกิจและการจัดหาพลังงาน นอกจากนี้ ความต้องการในการจัดทำรายงานผลการประกอบการปิโตรเลียมตามมาตรฐานของกรมเชื้อเพลิงธรรมชาติ ด้วยกระบวนการจัดเตรียมและตรวจสอบข้อมูลจากหลุมพัฒนาการผลิตจำนวนมากนี้เป็นความท้าทายที่สำคัญสำหรับบริษัทผู้ดำเนินงาน โดยเฉพาะเมื่อข้อมูลที่ได้รับมีความหลากหลายและอยู่ในรูปแบบที่ซับซ้อน เช่น ไฟล์ PDF ที่ประกอบด้วยข้อความและรูปภาพ

การจัดทำรายงานผลการประกอบการปิโตรเลียมต้องการการรวบรวมข้อมูลจากแหล่งต่าง ๆ และการตรวจสอบข้อมูลเพื่อให้เป็นไปตามมาตรฐานที่กำหนด ปัจจุบันกระบวนการนี้มีความซับซ้อนและใช้เวลานานเนื่องจากต้องจัดการกับข้อมูลจำนวนมากในหลายรูปแบบ ซึ่งมักจะทำให้เกิดข้อผิดพลาดและเพิ่มภาระงานให้กับเจ้าหน้าที่ที่ต้องตรวจสอบข้อมูลด้วยมือ

การวิจัยนี้จึงมุ่งเน้นในการพัฒนาและปรับปรุงกระบวนการอัตโนมัติ โดยการนำเทคโนโลยี Optical Character Recognition (OCR) และปัญญาประดิษฐ์ (AI) มาใช้ในการสกัดข้อมูลจากไฟล์ PDF เพื่อเพิ่มประสิทธิภาพในการจัดการข้อมูลการยังยั้งกรณีหลุมเจาะและการจัดทำรายงาน ลดข้อผิดพลาดในการประมวลผล และช่วยให้กระบวนการทำงานรวดเร็วและแม่นยำยิ่งขึ้น การพัฒนาระบบอัตโนมัติที่ใช้เทคโนโลยีเหล่านี้สามารถช่วยลดภาระงานที่ต้องใช้คนจำนวนมากและเพิ่มความแม่นยำในการจัดทำรายงาน ซึ่งจะส่งผลดีต่ออุตสาหกรรมปิโตรเลียมและช่วยให้การดำเนินงานมีประสิทธิภาพสูงขึ้น

วัตถุประสงค์ของงานวิจัย

1. เพื่อพัฒนาระบบอัตโนมัติ: สร้างระบบอัตโนมัติในการแยกข้อมูลส่วนหัวของการยังยั้งกรณีหลุมเจาะ (well log header) จากไฟล์ PDF ให้เป็นข้อความในรูปแบบตาราง Excel เพื่อการจัดการข้อมูลที่รวดเร็วและมีประสิทธิภาพ
2. เพื่อเพิ่มความรวดเร็วในการตรวจสอบ: ปรับปรุงกระบวนการตรวจสอบข้อมูลให้สะดวกและรวดเร็วขึ้น
3. เพื่อลดข้อผิดพลาด: ลดข้อผิดพลาดจากการทำงานด้วยมือและการแก้ไขข้อมูลผิดพลาดที่อาจส่งผลกระทบต่อภาพลักษณ์และความน่าเชื่อถือของบริษัท

ขอบเขตงานวิจัย

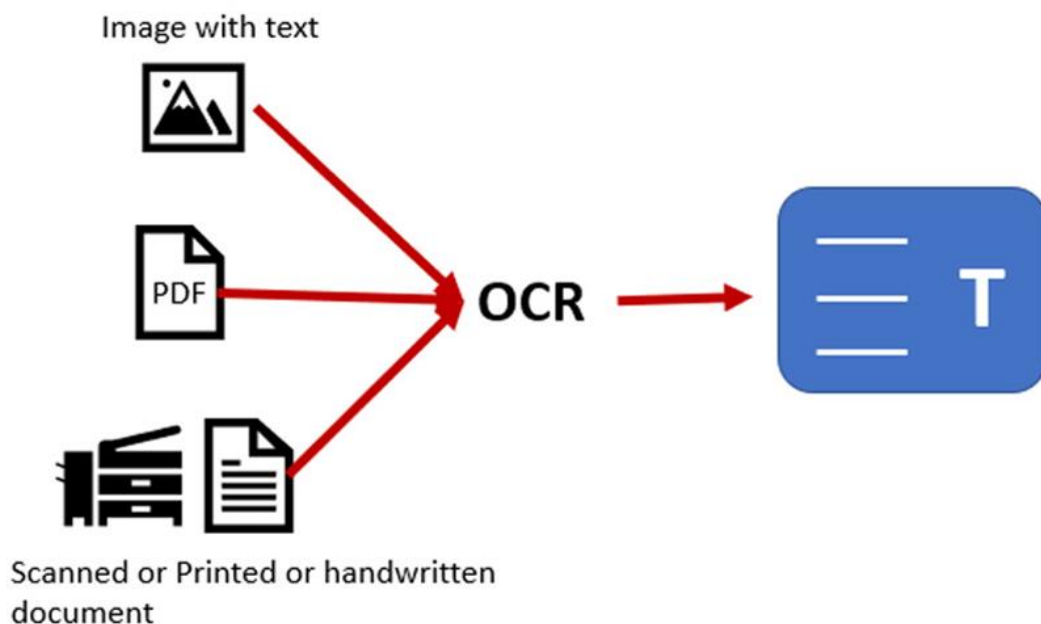
1. จัดการไฟล์ PDF: สกัดข้อมูลจากส่วนหัวของการยังยั้งกรณีหลุมเจาะ (well log header) ในไฟล์ PDF เช่น ชื่อหลุมเจาะ, วันที่, บริษัท, ความลึก, พิกัด และมาตราส่วน
2. ใช้ OCR Library: ใช้ Microsoft OCR Library ผ่าน AI Builder
3. พัฒนา AI Model: สร้างและปรับปรุง AI Model ผ่าน AI Builder
4. สร้างระบบอัตโนมัติ: ใช้ Power Automate เพื่อจัดการและประมวลผลข้อมูลในรูปแบบตาราง Excel

ทฤษฎี และผลงานที่เกี่ยวข้อง

งานวิจัยนี้มีวัตถุประสงค์ในการพัฒนาระบบอัตโนมัติสำหรับการแยกข้อมูลส่วนหัวของการยังยั้งกรณีหลุมเจาะ (well log header) เพื่อเพิ่มความสะดวกและรวดเร็วในการตรวจสอบความถูกต้องของข้อมูล รวมถึงลดข้อผิดพลาดในการทำงาน โดยใช้เทคโนโลยีการรู้จำอักขระด้วยแสง (OCR) หรือ Optical Character Recognition เพื่อแปลงข้อมูลจากรูปแบบอนาล็อกเป็นดิจิทัล และนำกระบวนการทำงานแบบอัตโนมัติมาประยุกต์ใช้เพื่อเพิ่มประสิทธิภาพในการจัดการข้อมูลและเอกสาร

1. การรู้จำอักขระด้วยแสง (OCR)

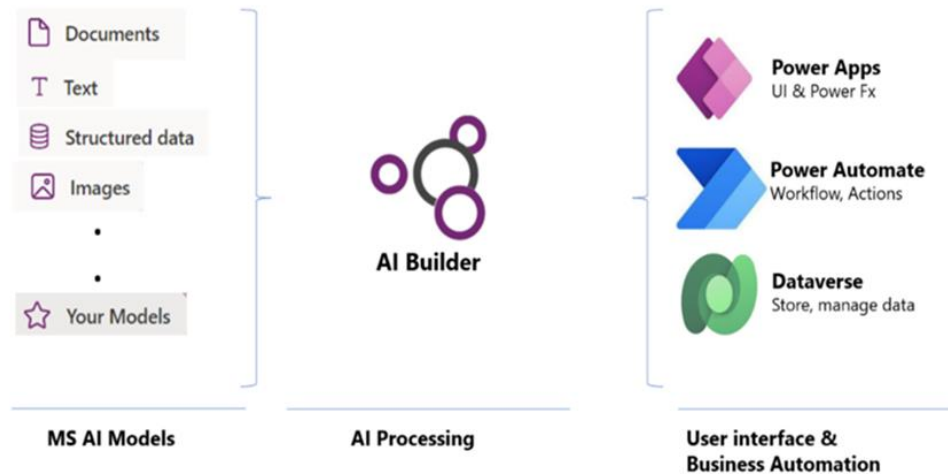
OCR ย่อมาจาก Optical Character Recognition เป็นเทคโนโลยีที่ช่วยให้ระบบคอมพิวเตอร์ อ่านข้อความจากรูปภาพ ออกมาอยู่ในรูปแบบของข้อความ ชุดของตัวอักษรที่สามารถนำไปจัดเก็บเพื่อสืบค้น เรียกข้อมูล ทั้งนี้ข้อมูลที่อ่านจากเอกสารอาจจะไม่สามารถอ่านได้ 100 % เนื่องจากโมเดลที่ใช้ สภาพของไฟล์เอกสาร ความคมชัด รายละเอียด ซึ่งมีผลต่อการอ่านข้อความ



ภาพที่ 1 ขั้นตอนการทำงานของเทคโนโลยี การรู้จำอักขระด้วยแสง (OCR)

2. AI Builder

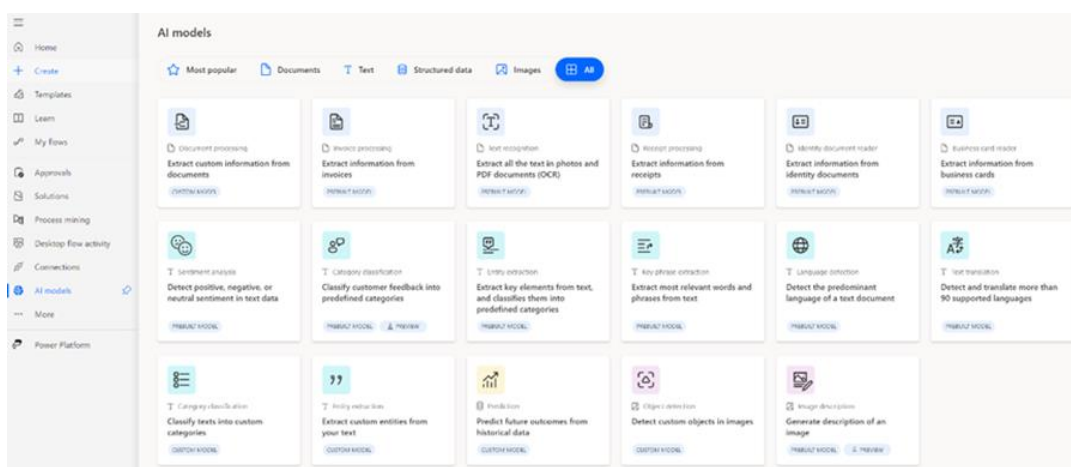
AI Builder เป็นเครื่องมือตัวหนึ่งของ Microsoft Power Platform ในการสร้างและปรับปรุง AI Model ที่มีการประมวลผลอัตโนมัติ เป็นส่วนหนึ่งของโซลูชันในการพัฒนาแอปพลิเคชันหรือระบบอัตโนมัติ ด้วย Power Apps หรือ Power Automate ประโยชน์ของ AI Builder คือช่วยให้ผู้ใช้สามารถสร้างโมเดล AI ที่สามารถเข้าใจและปรับปรุงประสิทธิภาพการทำงานของแอปพลิเคชัน หรือระบบได้อย่างง่ายดาย โดยไม่ต้องมีความรู้เฉพาะทางในด้าน AI



ภาพที่ 2 สถาปัตยกรรมของ AI Builder

3. AI Model

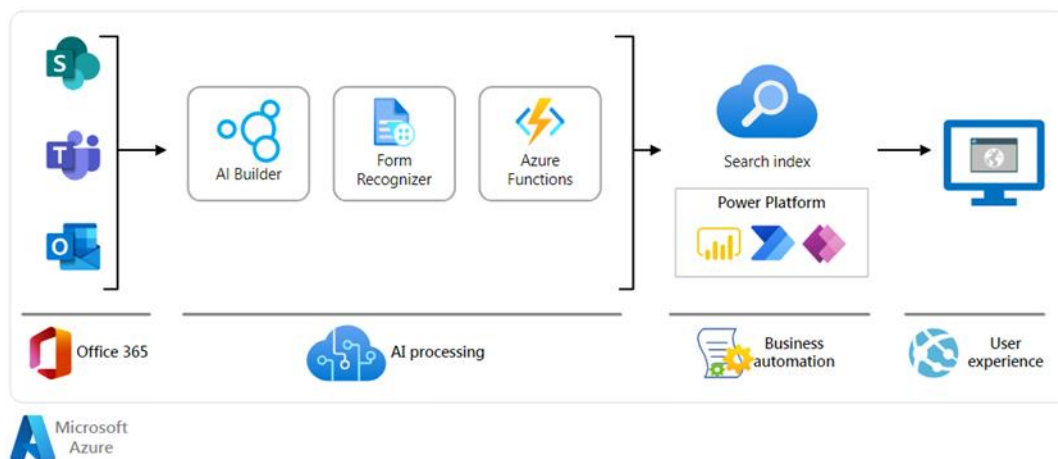
โมเดล AI ช่วยสร้างใช้INPUTของมนุษย์เพื่อเรียนรู้และปรับปรุงการคาดการณ์ ความสามารถของโมเดลเหล่านี้ในการสนับสนุนการตัดสินใจของมนุษย์คือการปรับใช้ที่สำคัญและบางครั้งอาจถูกมองข้าม การใช้งานที่เป็นไปได้ ได้แก่ การวินิจฉัยทางคลินิก ระบบสนับสนุนการตัดสินใจ และการวิเคราะห์ ความสามารถอีกอย่างหนึ่งคือการพัฒนาแอปพลิเคชัน AI แบบใหม่โดยการปรับแต่งโมเดลพื้นฐานที่มีอยู่อย่างละเอียด ได้อย่างสะดวกรวดเร็ว



ภาพที่ 3 ประเภทของแบบจำลอง AI (Microsoft Learn)

4. Power Automate

Power Automate เป็นซอฟต์แวร์ RPA ของ Microsoft ที่ใช้แพลตฟอร์ม Low-Code สำหรับสร้างและพัฒนา BOT บนคลาวด์ ทำงานร่วมกับ Microsoft 365, API, และ AI Builder ได้อย่างราบรื่น ช่วยให้ผู้ใช้เริ่มต้นได้ง่ายและฟรี โดยสามารถสร้างงานอัตโนมัติได้ในไม่กี่นาที เป็นส่วนหนึ่งของ Microsoft Power Platform ซึ่งรวมถึง Power Apps, Power BI, และ Power Virtual Agents มาพร้อมกับเทมเพลตจำนวนมากที่ช่วยให้ผู้ใช้ที่ไม่ต้องการเขียนโค้ดสามารถพัฒนา Workflow ได้อย่างรวดเร็ว ลดข้อผิดพลาดจากการทำงานด้วยมือและเพิ่มประสิทธิภาพในการทำงาน



ภาพที่ 4 สถาปัตยกรรมของ การสกัดข้อความจากวัตถุ โดยใช้ Power Automate และ AI Builder
(Microsoft Learn)

5. งานวิจัยที่เกี่ยวข้อง

Norbert Gal, Vasile Stoicu-Tivadar, Emanuela Gal-Nadasan and Anca Dinu “Robotic Process Automation Based Data Extraction from Handwritten Medical Forms” (2023) งานวิจัยนี้นำเสนอการสร้างระบบอัตโนมัติสำหรับการดึงข้อมูลจากแบบฟอร์มทางการแพทย์ที่เขียนด้วยลายมือ โดยใช้ RPA และ OCR. ระบบนี้ใช้ UiPath สำหรับการสร้างหุ่นยนต์ และ Google Cloud Vision OCR สำหรับการแปลงข้อมูลให้อยู่ในรูปแบบดิจิทัลผ่านไฟล์ DOM (Digital Object Model). เนื่องจากแบบฟอร์มมีรูปแบบที่แตกต่างกัน, จำเป็นต้องสร้างเทมเพลตเฉพาะและฝึกฝน AI เพื่อรองรับลายมือที่หลากหลาย. ปัจจัยที่ส่งผลต่อความแม่นยำคือรูปแบบลายมือที่กรอกฟอร์ม, โดยลายมือที่เขียนเป็นตัวพิมพ์ใหญ่จะมีความแม่นยำสูงกว่า

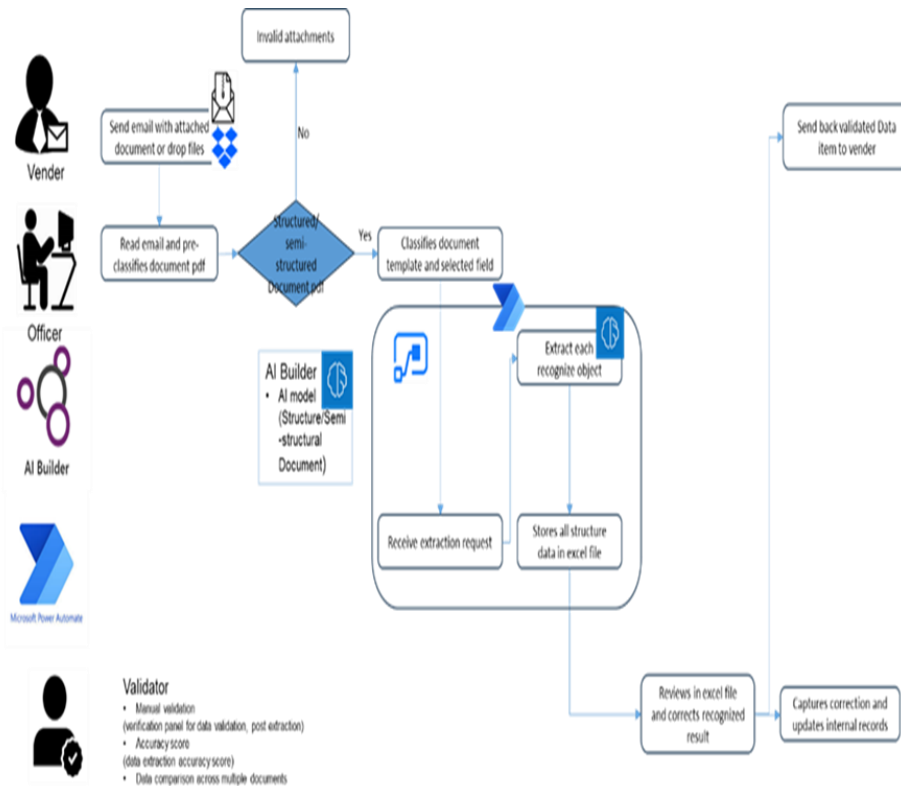
Nur Ashikin Rohaime; Nur Idora Abdul Razak; Norashikin M. Thamrin and Chua Wei Shyan, “Integrated Invoicing Solution: A Robotic Process Automation with AI and OCR Approach” : (2022) งานวิจัยนี้เสนอระบบอัตโนมัติสำหรับการสกัดข้อมูลจากใบแจ้งหนี้ที่จัดเก็บในรูปแบบไฟล์ CSV โดยใช้ TagUI (เครื่องมือ RPA โอเพ่นซอร์ส) และ PyTesseract (เครื่องมือ OCR) โซลูชันนี้มีความแม่นยำ 100% และใช้เวลาประมวลผลเร็วขึ้น 80% เมื่อเปรียบเทียบกับการใช้ IQBot (โซลูชันเชิงพาณิชย์) ที่ใช้ Naive Bayes ML algorithm และ natural language processing (NLP). งานวิจัยแสดงให้เห็นว่าโซลูชันโอเพ่นซอร์สมีประสิทธิภาพใกล้เคียงหรือดีกว่าในแง่ของความแม่นยำและความเร็วในการประมวลผล และมีความยืดหยุ่นและราคาต่ำกว่าเมื่อเปรียบเทียบกับโซลูชันเชิงพาณิชย์

ระเบียบวิธีวิจัย

การศึกษาวิจัยครั้งนี้ เป็นการนำเสนอระบบอัตโนมัติสำหรับการสกัดข้อมูลส่วนหัวของการยังธรณีหลุมเจาะ ด้วยเทคโนโลยีการรู้จำอักขระด้วยแสง (OCR) โดยผ่านการเทรน จาก AI builder เป็น AI model และใช้ Power Automate เป็นตัวสร้าง ระบบ RPA ที่พัฒนามาจาก Microsoft โดยมีแนวทางการวิจัยดังนี้

1. ระบบแนวคิด (System Flow)

แนวทางการสร้างระบบอัตโนมัติมีขั้นตอนดังภาพที่ 5



ภาพที่ 5 แผนผังการสร้างระบบอัตโนมัติ

2. การรวบรวมข้อมูล (Data Collection)

ในกระบวนการเก็บรวบรวมข้อมูล ข้อมูลการยังยธณิหลุมเจาะได้รับมาจาก ผู้ให้บริการผ่านทาง อีเมล โดยแนบชิปไฟล์ หรือ ส่ง ลิงค์ drop box เพื่อให้เข้าไป download เอกสาร โดย ผู้ปฏิบัติงานผู้เชี่ยวชาญ หรือผู้เกี่ยวข้องจะเป็นคนได้รับและ download เอกสารเก็บไว้

3. การทำความเข้าใจข้อมูล (Data Understanding)

จากข้อมูลที่ได้รับมา มีไฟล์ หลาขรูละแบบ ที่ต้องใช้โปรแกรมเฉพาะในการเปิดอ่าน ในการวิจัยครั้งนี้ เลือกลมาเฉพาะ Pdf ไฟล์ ที่มีข้อมูล ส่วนหัวของการยังยธณิหลุมเจาะ มาทำการสกัดข้อมูล

4. การเตรียมข้อมูล (Data Preparation)

จัดแยกรูปแบบเอกสาร ข้อมูลส่วนหัวการยังยธณิหลุมเจาะ ได้ทั้งหมด 8 รูปแบบ(template) แต่ละรูปแบบ (template) มีเขตข้อมูล (field) ที่ต้องการสกัดออกมา ตั้งแต่ 24 ถึง 35 เขตข้อมูล (field) ดังภาพที่ 6

Field	Mud log			LWD			Wireline	
	ML_SL	ML_HL	ML_WE	LWD_SL	LWD_HL	LWD_WE	RFT_SL	RFT_HL
1 vendor	vendor	vendor	vendor	vendor	vendor	vendor	vendor	vendor
2 well name	well name	well name	well name	logscale_1	logscale	logscale	company_1	company_1
3 start depth	start depth	start depth	start depth	logscale_2			well name_1	well name_1
4 total depth	total depth	total depth	total depth	logtype	logtype	logtype	field_1	field_1
5 scale_1	scale_1	scale_1	scale_1	company	company_1	company_1	rig_1	rig_1
6 scale_2	scale_2	scale_2	scale_2	company_2	company_2	company_2	country_1	country_1
7 X	X	X	X	well name	well name_1	well name_1	WL tool type	WL tool type
8 Y	Y	Y	Y	well name_2	well name_2	well name_2	run no_1	run no_1
9 lat	lat	lat	lat	field	field_1	field_1	rig_2	rig_2
10 long	long	long	long	field_2	field_2	field_2	field_2	field_2
11 RKB-MSL	RKB-MSL	RKB-MSL	RKB-MSL	Block	Block_1	Block_1	well name_2	well name_2
12 Water depth	Water depth	Water depth	Water depth	Block_2	Block_2	Block_2	company_2	company_2
13 9-5/8 csg.	9-5/8 csg.	9-5/8 csg.	9-5/8 csg.	rig_1	rig_1	rig_1	X	X
14 7 csg.	7 csg.	7 csg.	7 csg.	rig_2	rig_2	rig_2	Y	Y
15 2-7/8 csg.	2-7/8 csg.	2-7/8 csg.	2-7/8 csg.	hole size	hole size	hole size	KB	KB
16 mud type_ surface section	mud type_ surface section	mud type_ surface section	mud type_ surface section	country	country_1	country_1	country_2	country_2
17 Depth mud type_ surface section	Depth mud type_ surface section	Depth mud type_ surface section	Depth mud type_ surface section	country_2	country_2	country_2	lat	lat
18 mud type_ intermediate section	mud type_ intermediate section	mud type_ intermediate section	mud type_ intermediate section	lat	lat_1	lat_1	long	long
19 Depth mud type_ intermediate section	Depth mud type_ intermediate section	Depth mud type_ intermediate section	Depth mud type_ intermediate section	lat_2	lat_2	lat_2	logging date	logging date
20 mud type_ production section	mud type_ production section	mud type_ production section	mud type_ production section	long	long_1	long_1	run no_2	run no_2
21 Depth mud type_ production section	Depth mud type_ production section	Depth mud type_ production section	Depth mud type_ production section	long_2	long_2	long_2	Fluid type	Fluid type
22 12-3/4"	12-3/4"	12-3/4"	12-3/4"	X	X	X	max. record temp	max. record temp
23 8-1/2"	8-1/2"	8-1/2"	8-1/2"	Y	Y	Y	Block_1	Block_1
24 6-1/8"	6-1/8"	6-1/8"	6-1/8"	start date logged	start date logged	start date logged	Block_2	Block_2
25 field	field	field	field	complete date logged	complete date logged	complete date logged		
26 block	block	block	block	start depth logged	start depth logged	start depth logged		
27 country	country	country	country	total depth logged	total depth logged	total depth logged		
28 well type	well type	well type	well type	spud date	spud date	spud date		
29 contractor	contractor	contractor	contractor		total depth MD	total depth		
30 rig name	rig name	rig name	rig name		total depth TVD	total depth		
31 rig type	rig type	rig type	rig type		KB	KB		
32 spud date	spud date	spud date	spud date	rig type				
33 end of operations	end of operations	end of operations	end of operations					
34 TD reach	TD reach	TD reach	TD reach					
35 TD depth	TD depth	TD depth	TD depth					
36 TD depth_TVD-MSL								

ภาพที่ 6 ตารางขอบเขตข้อมูลที่ต้องการสกัดจาก ข้อมูลส่วนหัวการยิงธรณีหลุมเจาะ

5. การสร้างแบบจำลอง (Modeling)

1. สร้างและฝึกสอน โมเดล:

- เลือกประเภทโมเดล และปรับแต่งตามความต้องการ.
- เตรียมข้อมูล โดยใช้ไฟล์ตัวอย่างหลากหลายเพื่อฝึกสอนโมเดล.
- ปรับแต่ง การ mapping เพื่อให้โมเดลเรียนรู้ได้แม่นยำ.

2. ทดสอบและใช้งาน:

- Quick Test โมเดลกับไฟล์ทดสอบ.
- Publish โมเดลเพื่อใช้งาน.
- ใช้ใน Workflow ผ่าน Power Automate.

3. ประเมินประสิทธิภาพ:

- สร้างชุดข้อมูลทดสอบ ที่หลากหลาย.
- ประมวลผลเอกสาร เพื่อตรวจสอบความแม่นยำของโมเดล.

6. การสกัดข้อมูล (Data Extraction)

ใช้กระบวนการ OCR และ RPA ในการการสกัดข้อมูลจากเอกสารข้อมูลส่วนหัวการหยั่งธรณีหลุมเจาะแบบอัตโนมัติ ผ่าน Power Automate เป็นเครื่องมือหลักในการทำงานแบบ Automation และความสามารถในการเชื่อมต่อระบบต่างๆ เข้ากับบริการและ Application ของ Microsoft โดยเราจะนำเครื่องมือนี้มาใช้ integration ระหว่าง AI Builder และ flow การทำงาน เพื่อนำข้อมูลที่สกัดได้ มาแสดงในรูปแบบ excel file

ผลการศึกษา

สารนิพนธ์นี้ได้นำวิธีการรู้จำอักขระด้วยแสง (OCR) มาใช้ในการสกัดข้อมูลส่วนหัวของการหยั่งธรณีหลุมเจาะด้วยเทคโนโลยีปัญญาประดิษฐ์ สรุปได้ดังนี้

1. ระบบสามารถใช้เรียกไฟล์เพื่อตรวจสอบข้อมูลหัวหลุมเจาะเพื่อความต้องการของข้อมูลที่ปรากฏบนหัวเอกสาร
2. การใช้เทคนิค OCR ช่วยสกัดข้อมูลจากเอกสารส่วนหัวการหยั่งธรณีหลุมเจาะสามารถสกัดข้อมูลได้ดีมาก
3. การนำ OCR ตรวจสอบคำที่แยกมาแล้วเพื่อช่วยตรวจสอบข้อผิดพลาด โดยในการวิจัยจากตัวอย่างพบว่าตรวจสอบและเห็นข้อผิดพลาดได้ง่ายขึ้น
4. การใช้ Power Automate ช่วยลดการทำงานของเจ้าหน้าที่ปฏิบัติการในการเตรียมและตรวจสอบข้อมูล

ผลการวัดประสิทธิภาพความถูกต้องของ AI Model

ข้อมูลที่ใช้ในการฝึกอบรมจำนวน 139 ชุดข้อมูล จากข้อมูลทั้งหมด 199 ชุดข้อมูลแบ่งเป็น 8 แบบฟอร์ม ได้ดังแสดงในตาราง สรุปผลการเปรียบเทียบประสิทธิภาพ ตารางที่ 1

ตารางที่ 1 สรุปผลการเปรียบเทียบประสิทธิภาพความถูกต้อง 8 แบบฟอร์ม

AI Model	TRAIN			TEST	Total
	%AUC	Field/Example	Example	Example	Example
ML_SL	98	35	6	3	9
ML_HL	96	35	18	8	26
ML_WE	97	35	9	4	13
LWD_SL	99	22	24	10	34
LWD_HL	99	30	24	10	34
LWD_WE	98	25	24	10	34
RFT_SL	99	24	18	8	26
RFT_HL	97	24	16	7	23
Total		230	139	60	199

ผลการสกัดข้อมูลโดยใช้ AI Model ร่วมกับ Power Automate

ข้อมูลที่ใช้ในการสกัดข้อมูลจำนวน 60 ชุดข้อมูล ได้ดังแสดงในตาราง แสดงผลการสกัดข้อมูล ส่วนหัวการหยั่งธรณีหลุมเจาะ ภาพที่ 7 และ 8

Tool type	Vender	Start depth	Total depth	Scale_1	Scale_2	X	Y	Lat	Long	KB	Water depth	9-5/8" csg.	7" csg.
Mud Log	G eoservices	109.73		20981: 500			779 572.551	1 036 227.091	101° 32' 44.570" E	:		303.84	
Mud Log	G eoservices	73.46		1504.46 1: 500			779 572.551	1 036 227.091	101° 32' 44.570" E		263.86	261.05	
Mud Log	G eoservices	109.73		20981: 500			779 572.551	1 036 227.091	101° 32' 44.570" E	:		303.84	
Mud Log	G eoservices	73.46		1504.46 1: 500			779 572.551	1 036 227.091	101° 32' 44.570" E		263.86	261.05	
Mud Log	G eoservices	103.76 mTD TO: 3694.00 mTD SCALE: 1: 500					779 572.551	1 036 227.091	009° 21' 56.978" N 101° 32' 44.570" E	36.27	73.46	308.17	2026.01
Mud Log	G eoservices	109.73		34131: 500			779 567.759	1 036 225.740	009° 21' 56.935" N 101° 32' 44.412" E	36.27	73.46	304.97	2171.03
Mud Log	G eoservices	73.46		26651: 500			779 572.551	1 036 227.091	009° 21' 56.978" N 101° 32' 44.570" E	36.27	73.46	268	1496.79
Mud Log	G eoservices	109.73		33411: 500			779 568.589	1 036 226.980	009° 21' 56.975" N 101° 32' 44.440" E	36.27	73.46	307.69	1998.23
Mud Log	G eoservices	73.46		2590.311: 500			779 568.589	1 036 226.980	009° 21' 56.975" N 101° 32' 44.440" E	36.27	73.46	265.15	1556.82
Mud Log	G eoservices	109.73		30321: 500			779 572.741	1 036 225.070	009° 21' 56.912" N 101° 32' 44.575" N	36.27	73.46	308.09	1721.75
Mud Log	G eoservices	73.46		2535.531: 500			779 567.759	1 036 225.740	009° 21' 56.935" N 101° 32' 44.412" E	36.27	73.46	264.43	1552.33
Mud Log	G eoservices	73.46		2598.851: 500			779 572.741	1 036 225.070	009° 21' 56.912" N 101° 32' 44.575" N	36.27	73.46	263.4	1582.52
Mud Log	G eoservices	109.73		36281: 500			779 573.832	1 036 225.910	009° 21' 56.939" N 101° 32' 44.611" E	36.27	73.46	305.81	2147.44
Mud Log	G eoservices	73.46		2412.21: 500			779 573.832	1 036 225.910	009° 21' 56.939" N 101° 32' 44.611" E	36.27	73.46	264.22	1284.69
Mud Log	G eoservices	73.46		2598.851: 500			779 572.741	1 036 225.070	009° 21' 56.912" N 101° 32' 44.575" N	36.27	73.46	263.4	1582.52

ภาพที่ 7 ตัวอย่างตารางแสดงผลการสกัดข้อมูลส่วนหัวการหยั่งธรณีหลุมเจาะ 1

Tool type	Vender	Log scale_1	Log scale_2	Log type	Field	Rig_1	Rig_2	Hole size	Country	Lat	Long	Block	Rig type
LWD	slb	1:200	1:200	MD	TRAT G1/61	Idun	Idun	6.125in	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:200	1:200	TVD-MSL	TRAT G1/61	Idun	Idun	6.125in	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:500	1:500	MD	TRAT G1/61	Idun	Idun	6.125in	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:500	1:500	TVD-MSL	TRAT G1/61	Idun	Idun	6.125in	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:500	1:500	TVD-MSL	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:200	1:200	MD	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:500	1:500	MD	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:200	1:200	TVD-MSL	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.912" N	101° 32' 44.575" E	G1/61	Jack Up
LWD	slb	1:200	1:200	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:500	1:500	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:200	1:200	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:500	1:500	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:500	1:500	TVD-MSL	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:200	1:200	MD	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:200	1:200	TVD-MSL	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:500	1:500	MD	TRAT G1/61	Idun	Idun	8.5	Thailand	9° 21' 56.978" N	101° 32' 44.57" E	G1/61	Jack Up
LWD	slb	1:500	1:500	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.975" N	101° 32' 44.44" E	G1/61	Jack Up
LWD	slb	1:500	1:500	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.975" N	101° 32' 44.44" E	G1/61	Jack Up
LWD	slb	1:200	1:200	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.975" N	101° 32' 44.44" E	G1/61	Jack Up
LWD	slb	1:200	1:200	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.975" N	101° 32' 44.44" E	G1/61	Jack Up
LWD	slb	1:200	1:200	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.935" N	101° 32' 44.412" E	G1/61	Jack Up
LWD	slb	1:500	1:500	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.935" N	101° 32' 44.412" E	G1/61	Jack Up
LWD	slb	1:200	1:200	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.935" N	101° 32' 44.412" E	G1/61	Jack Up
LWD	slb	1:500	1:500	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.935" N	101° 32' 44.412" E	G1/61	Jack Up
LWD	slb	1:500	1:500	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up
LWD	slb	1:200	1:200	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up
LWD	slb	1:500	1:500	MD	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up
LWD	slb	1:200	1:200	TVD-MSL	TRAT G1/61	Idun	Idun	6.125	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up
LWD	slb	1:200	1:200	Subsea	G1/61 TRAT	8.5 Borr	Idun	8.5	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up
LWD	slb	1:500	1:500	Subsea	G1/61 TRAT	8.5 Borr	Idun	8.5	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up
LWD	slb	1:200	1:200	Depth	G1/61 TRAT	8.5 Borr	Idun	8.5	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up
LWD	slb	1:500	1:500	Depth	G1/61 TRAT	8.5 Borr	Idun	8.5	Thailand	9° 21' 56.939" N	101° 32' 44.611" E	G1/61	Jack Up

ภาพที่ 8 ตัวอย่างตารางแสดงผลการสกัดข้อมูล ส่วนหัวการหยั่งธรณีหลุมเจาะ 2

สรุปและอภิปรายผลการวิจัย

1. ได้ระบบอัตโนมัติ ที่สามารถสกัดข้อมูลส่วนหัวการหยั่งธรณีหลุมเจาะ เพื่อ ตรวจสอบความถูกต้องก่อนส่งข้อมูล ให้กรมเชื้อเพลิงธรรมชาติ
2. การใช้เทคนิค OCR (Optical Character Recognition) ช่วยให้เราสามารถสกัดข้อมูลจากเอกสาร PDF ที่มีรูปแบบที่กำหนดไว้อย่างชัดเจนได้อย่างมีประสิทธิภาพและเป็นไปตามข้อกำหนดที่กำหนดไว้อย่างไรก็ตาม การใช้เทคนิค OCR นั้นต้องมีการกำหนดรูปแบบเอกสาร (Structure document) อย่างชัดเจนเพื่อให้เทคโนโลยีสามารถทำงานได้อย่างมีประสิทธิภาพ
3. การนำผลการอ่านของ AI model เข้ามาช่วยให้เจ้าหน้าที่สามารถโฟกัสเฉพาะในข้อมูลที่มีความผิดพลาดหรือข้อมูลที่ต้องการการตรวจสอบเพิ่มเติมได้อย่างมีประสิทธิภาพ เช่นการระบุข้อมูลที่คล้ายคลึงกันหรือการค้นหาข้อมูลที่หายไป
4. การใช้เทคโนโลยีเหล่านี้ช่วยลดเวลาในการทำงานของเจ้าหน้าที่และเพิ่มความถูกต้องในการปฏิบัติงาน เนื่องจากการประมวลผลข้อมูลอัตโนมัติและลดความเสี่ยงของข้อผิดพลาดมากขึ้น

ข้อเสนอแนะ

1. ควรพัฒนาโมเดลให้สามารถทำงานที่หลากหลายรูปแบบมากขึ้น ทำให้สามารถนำโมเดลประยุกต์ใช้กับข้อมูล การหยังรรณิหลุมเจาะรูปแบบอื่นได้
2. ควรมีการกำหนด มาตรฐานรูปแบบของ แบบฟอร์ม ข้อมูลส่วนหัวการหยังรรณิหลุมเจาะ ให้คงที่ เพื่อให้ข้อผิดพลาดลดน้อยลง
3. ควรมีการเพิ่มเอกสารรูปแบบใหม่ เพิ่มเติมเพื่อช่วยเพิ่มประสิทธิภาพของการสร้างแบบจำลองมากขึ้น และอัปเดตโมเดล ทุกครั้งที่มีการเปลี่ยนแปลงรูปแบบเอกสาร

บรรณานุกรม

- Norbert Gal-Nadasan, Vasile Stoicu-Tivadar, Emanuela Gal-Nadasan, and Anca Raluca Dinu. (2023). Robotic Process Automation Based Data Extraction from Handwritten Medical Forms. *EFMI-STC 2023*, pp 68-72
- Nur Ashikin Rohaime, Nur Idora Abdul Razak, Norashikin M. Thamrin and Chua Wei Shyan. (2021). Integrated Invoicing Solution: A Robotic Process Automation with AI and OCR Approach. *IEEE*, pp 30-33.