

Enhancing Thai Rice Query Assistance through a Knowledge-Driven Approach Using GraphRAG

Gampanut Soontontam¹, Tinnaphob Dindam², Adisorn Kheaksong³, Kanabadee Srisomboon⁴, and Parinya Sanguansat^{5*}

^{1,2,3,5}Faculty of Engineering and Technology, Panyapiwat Institute of Management, Nonthaburi, Thailand

⁴Department of Electrical and Computer Engineering, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

E-mail: 6652300168@stu.pim.ac.th, tinnaphobdin@pim.ac.th, adisornkhe@pim.ac.th, kanabadee.s@eng.kmutnb.ac.th, parinyasan@pim.ac.th*

Received: February 28, 2025 / Revised: April 4, 2025 / Accepted: May 19, 2025

Abstract—Thai rice farmers face significant challenges accessing timely and accurate information for crucial decisions regarding variety selection, soil management, and adapting to climate change. While Retrieval-Augmented Generation (RAG) systems aim to provide information, traditional RAG often struggles with complex queries requiring interconnected knowledge and can yield generic or less relevant answers in specialized domains like agriculture due to its reliance on the semantic similarity of isolated text chunks. This paper introduces and evaluates GraphRAG, a knowledge graph-enhanced RAG approach, designed specifically to overcome these limitations and improve query assistance for Thai rice cultivation. The methodology involves constructing a knowledge graph from key Thai rice farming documents and integrating it with a large language model to provide context-aware responses, comparing its performance against a traditional RAG baseline. Results demonstrate GraphRAG's superior effectiveness; user preference tests showed participants favored GraphRAG responses (52.9%) significantly more than traditional RAG (35.3%), particularly for complex queries requiring nuanced understanding. Quantitatively, GraphRAG showcased enhanced efficiency, reducing the average query response time by nearly 3 times (from 1.43 seconds for RAG to 0.41 seconds) and decreasing memory usage by over 50% (from 457.42 KB for RAG to 213.09 KB). This study concludes that GraphRAG offers a valuable approach for enhancing information retrieval accuracy, contextual understanding, and system efficiency in specialized, low-resource agricultural domains, highlighting its significance for providing better decision support to farmers.

Index Terms—GraphRAG, LLM, Thai Rice, Knowledge Graph, Retrieval-Augmented Generation

I. INTRODUCTION

Rice (*Oryza sativa* L.) is the main food for more than half of the world's population. Mostly grown in tropical zones, rice is full of carbohydrates, proteins, and healthy compounds, which greatly help with various health benefits. Recent improvements in rice science, especially in Androgenesis, have made breeding methods better, but issues like dependence on the type of plant still exist. Besides its role in food security, rice's nutrition, which includes important vitamins and minerals, makes it a key part of many food items and health supplements [1]. Also, farming actions, like when planting is done, greatly affect how much rice is produced, highlighting how important it is to make growing methods as good as possible [2]. Using technology, such as machine learning to identify seed types, in rice farming is also causing big changes [3]. These points together show how rice is important in many ways for both farming and public health, which will be talked about more in the next parts. Growing rice is very important in Thailand, as both a major economic product and a cultural symbol. Thailand is one of the top countries selling rice worldwide, with types like Jasmine rice (*Oryza sativa* var. *Indica*) becoming known around the world for their special smell. However, rice farmers have many problems, including picking the best rice types, managing their soil, and adapting to climate change. Normal ways of sharing information, like agricultural extension services, are often slow and not very effective, which stops farmers from getting timely and useful information [4]. To fix these problems, ways that use knowledge and rely on Artificial Intelligence (AI) and Machine Learning (ML) are becoming more and more needed in farming [5].

The rise of Large Language Models (LLMs) has played a very important role in natural language processing, demonstrating remarkable capabilities in generating coherent text, answering complex questions, and even producing creative content. Their ability to learn from massive datasets has positioned them as powerful tools across diverse applications. However, despite their impressive general knowledge, LLMs inherently suffer from a critical limitation: a lack of up-to-date, specialized, and contextually relevant domain knowledge [6]. This deficiency becomes particularly pronounced when addressing tasks requiring precision and deep understanding within specific fields, especially in agriculture.

Traditional methods of fine-tuning LLMs on domain-specific datasets can be computationally expensive. Practically, Retrieval-Augmented Generation (RAG) systems offer good answers by mixing the ability to find documents with the language power of Large Language Models (LLMs) to create correct and relevant responses by retrieving relevant documents or information snippets at runtime [7]. However, normal RAG systems often have trouble with questions about specific topics because they are not very good at understanding complex relations between pieces of data and entities [8]. To overcome this problem, systems like GraphRAG use knowledge graphs to produce a graph-structured knowledge representation that depicts domain hierarchies and entity interactions [9]. The design of these knowledge-based solutions will be looked at closely throughout this paper.

The GraphRAG system is designed to solve these issues by including a knowledge graph. This organizes data into connected items and links, thus helping to give more thoughtful and helpful answers to questions [10]. In this paper, the Llama 3 8b model is utilized as the Large Language Model (LLM), and Neo4j is used as a graph representation, which has shown it is good at understanding and creating natural language. GraphRAG can make more precise and contextually suitable answers. Furthermore, we use LangChain to help the system's ability to take out useful data from PDF documents, adding to the knowledge graph and allowing for more complete handling of questions [11]. This paper introduces GraphRAG, a knowledge graph-based extension of Retrieval-Augmented Generation (RAG), designed to enhance the retrieval and generation of Thai rice-related information.

We compare GraphRAG with traditional RAG models in this specific task. The remainder of this paper is organized as follows: Section II reviews related work in knowledge graphs and RAG systems, Section III details the proposed GraphRAG methodology and experimental setup, Section IV presents and discusses the results of our comparative

analysis, and Section V concludes the paper with key findings and future research directions. The results demonstrate that GraphRAG not only outperforms standard RAG systems, particularly when handling complex queries and extracting useful insights from unstructured data sources, but also remains effective even in low-resource settings within the Thai rice domain.

II. RELATED WORK

Past research has demonstrated the value of incorporating knowledge graphs (KGs) to enhance information searching and Question-Answering (QA) systems across various domains. However, traditional QA systems and even standard Retrieval-Augmented Generation (RAG), which retrieves text chunks based on vector similarity to augment Large Language Models (LLMs), can face limitations in handling complex queries requiring deep contextual understanding or reasoning over interconnected information [8].

Recent advancements have led to the development of GraphRAG, an approach that specifically integrates KG construction and querying within the RAG framework to address these limitations [9]. GraphRAG has been explored in various contexts, such as improving accuracy in complex domains like healthcare support [12] or enhancing document analysis through structured knowledge extraction [13]. Unlike standard RAG, GraphRAG typically involves dynamically extracting entities and relationships from the source documents themselves to build a knowledge graph, which then provides richer, interconnected context for the LLM's response generation.

While GraphRAG offers a novel integration, other research has focused on leveraging pre-existing or separately constructed KGs. For example, the UCKG-Why-QA system demonstrated the utility of a cause-and-effect KG for diagnosing plant diseases, including rice sicknesses, showing improved accuracy for complex questions through visualization [15]. However, it encountered challenges with fine-grained farming details. Similarly, Xie *et al.* [16] utilized Neo4j and BERT-CRF to build a dedicated TCM knowledge graph (with over 2,200 entities and 5,100 relationships), enhancing specialized searches but facing limitations due to data availability.

Comparing these approaches, GraphRAG differs significantly. While UCKG-Why-QA and the TCM system rely on potentially pre-defined or separately curated KGs, GraphRAG often constructs its knowledge graph directly from the input text corpus used for retrieval. This tight integration aims to provide context that is highly relevant to the source documents and potentially offers greater scalability and adaptability compared to methods requiring

extensive manual ontology engineering or facing the data completeness issues noted by Xie *et al.* [16]. Furthermore, GraphRAG's emphasis on relationship traversal within the generated graph provides a mechanism for deeper reasoning compared to the semantic chunk retrieval of standard RAG. While other methods also exist to enhance RAG, such as sophisticated re-ranking algorithms applied after initial retrieval [14], this study focuses specifically on the potential of the integrated knowledge graph construction offered by GraphRAG.

Therefore, this research investigates the application and effectiveness of the GraphRAG approach within the specific, low-resource domain of Thai rice farming, comparing it against a standard RAG baseline to evaluate its benefits in addressing complex farmer queries.

III. METHODOLOGY

A. Knowledge Graph Visualization

To illustrate how different parts of the Thai rice system are linked and function together, we have constructed a knowledge graph. This graph is essentially a visual representation that utilizes nodes and relationships to organize information. In this paper, we built the knowledge graph as shown in Fig.1 from our data collection, which will be explained later. However, it is hard to view in detail because of its size. We show some of the chunks as follows:

Fig. 2 is the Thai rice industry knowledge graph; the nodes and edges in this graph are:

- Business People: The main part is the Thai Rice Business; the lines connect it to important groups that research Thai Farmers and groups that study rice selling.

- Countries: Thailand is shown along with countries that are its neighbors and also sell rice (Cambodia, Myanmar, Vietnam, India, and Pakistan). This shows who competes with whom and who trades with whom.

Types of Connections: The lines (with labels) show different kinds of connections. Here's an example of the connections:

- EXPORTS TO: This Shows where Thai rice is exported.

- COMPETITO: Points out which countries are competitors with Thailand in selling rice.

- COMPARED TO: Shows when Thai rice is being compared to or studied against something.

This graph helps us understand better how the Thai rice business is set up, how rice flows in trade, and who the competitors are.

Fig. 3 is the Thailand Agricultural Knowledge Graph. The nodes and edges in this graph are:

- GOVERNMENT AND TEAMS: The main part is the Agriculture Ministry, and the lines connect it

to important teams like the Agriculture Department, Trade Department, and the Bank for Agriculture. This shows the government teams that help and manage Thai farming.

- MAIN FARM PRODUCT & NATION: Thai Rice is the main product; it's connected to Thailand, showing rice is grown there. Thailand is in the middle, linked to all parts of its farming, showing it runs and helps farming overall.

Types of Connections: The lines (with labels) show different connections. Here are some examples:

- HAS BANK: Shows the Agriculture Ministry links to the Bank for Agriculture, like it's in charge or related.

- HAS AGENCY: Shows the Agriculture Ministry is in charge of Thai Rice.

- PRODUCES: Show what's made in Thailand, like Thai Rice is made in Thailand, and Farming makes food in Thailand.

This graph helps us see how Thai farming is set up, showing how government teams, offices, banks, and farm products are all linked in the country.

Fig. 4 is the Rice Processing Knowledge Graph; the nodes and edges in this graph are:

- RICE COMPONENTS: Milled Rice is the main focus, with related parts around it, such as Rice Husk, Rice Bran, Sticky Rice, and White Rice Seed. These show the different types of rice or materials used when processing rice. This describes the various rice materials involved.

- STEPS IN RICE MAKING: You can see Processing and Contamination. Processing is the way rice is changed from its seed to what we eat. Contamination refers to when rice becomes unclean or contains unwanted substances. These points out the important actions and possible problems during rice production.

Types of Connections: The lines with labels show different kinds of relationships:

- CONTAMINATION: Indicates Milled Rice and Sticky Rice can experience Contamination, meaning they can become impure.

- PROCESSING: This shows that Milled Rice is made through Processing, which changes it to White Rice Seed.

- CONTAINS: Suggests Professor CONTAINS information about Milled Rice, implying experts milled rice.

This graph gives a structured overview of rice processing, covering the rice materials, the steps involved, the knowledge sources, and potential quality issues, all in simple terms.

Fig. 5 is the Rice Disease Knowledge Graph; the nodes and edges in this graph are:

- RICE DISEASES: The main point is that Rice Diseases, Bakanae, Fusarium Moniliforme, and "Virus" are neighbors. These are examples of different sicknesses that can affect rice plants. This shows the types of diseases rice can get.

- CAUSES OF DISEASE: Fusarium Moniliforme and Virus. These are things that CAUSE or CONTRIBUTE TO the Rice Disease happening in rice plants. This tells us what makes rice become diseased.

- SPREADING DISEASE: There is an Insect Vector. This shows how a Virus can be transmitted by insects. Insects can move the “Virus” around to different rice plants. This shows how some rice diseases can spread.

Types of Connections: The lines with words explain the links between them:

- CAUSED BY: Shows that Rice Disease is CAUSED BY Fusarium Moniliforme—this fungus makes the disease happen.

- CAUSES: Shows that Fusarium Moniliforme CAUSES Bakanae—this fungus leads to another disease called Bakanae.

- CONTRIBUTES TO: Shows that Virus

- CONTRIBUTES TO Rice Disease—Viruses are also part of the rice disease problem.

This graph helps to understand rice diseases, what causes them, how they spread, and the different kinds that can affect rice crops.

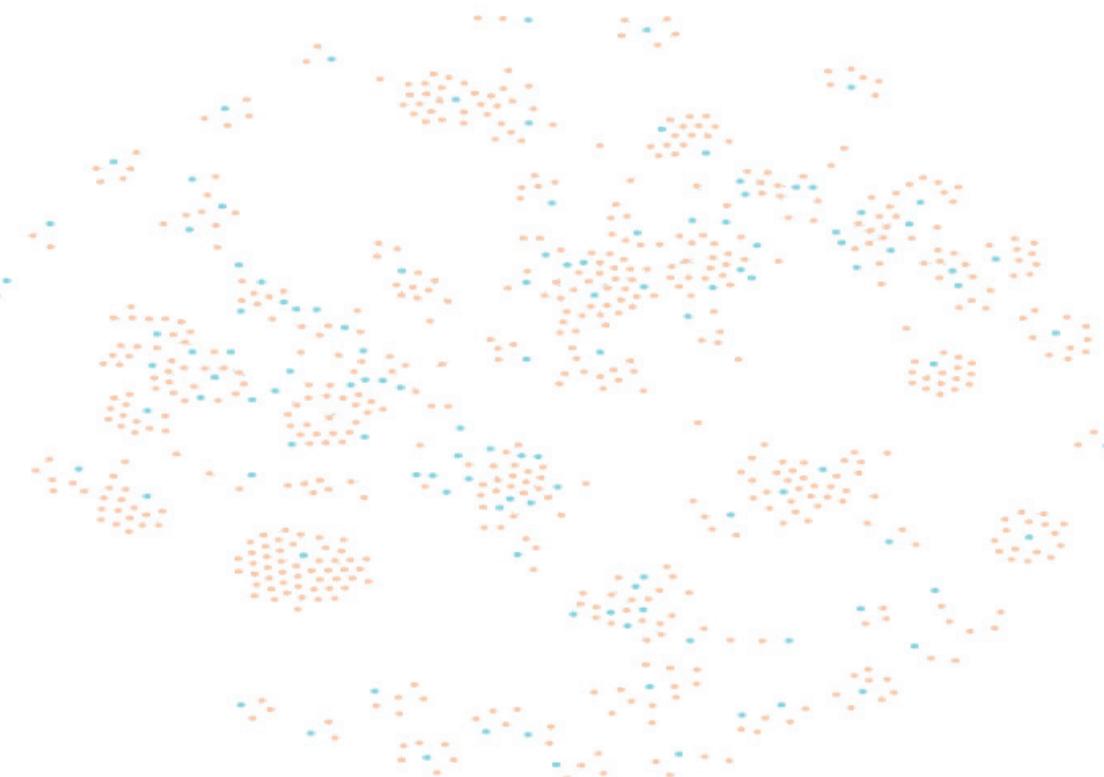


Fig. 1. The knowledge graph for Thai rice that includes hundreds of chunks with nodes and edges visualization using neo4j graph database

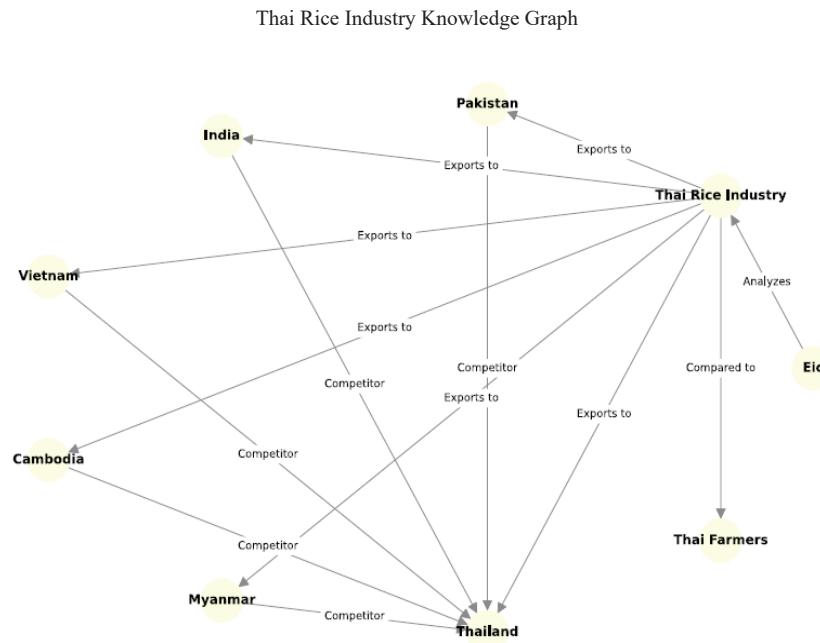


Fig. 2. Thai rice industry knowledge graph. It shows the important people involved, who sells rice to whom, how they compete, and compares them with other countries in Southeast Asia. Dots in the picture are things like Thailand, Thai Farmers, and Vietnam. Lines show connections like EXPORTS TO, COMPETITOR, COMPARED TO.

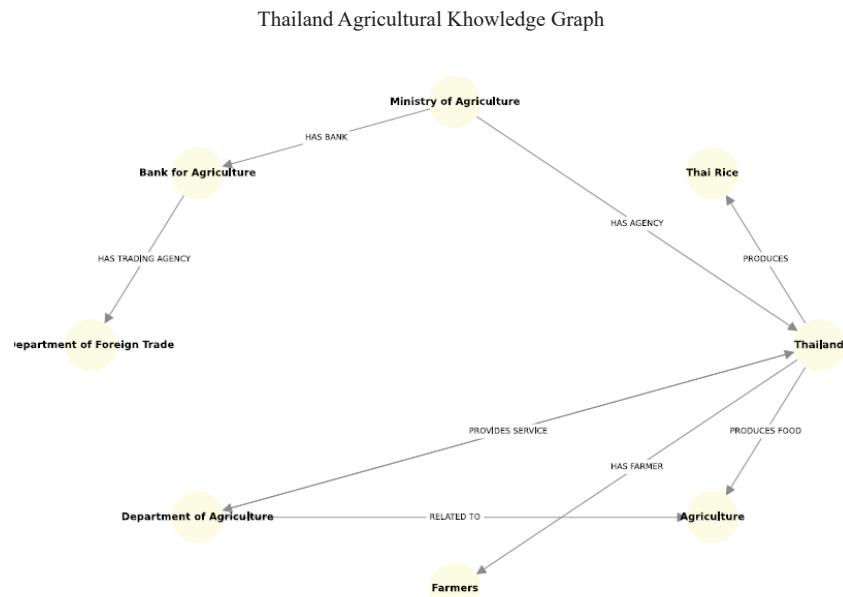


Fig. 3. Thailand Agricultural Knowledge Graph. It depicts key entities within the Thai agricultural domain, such as government ministries, agricultural sectors (like Thai Rice), and stakeholders (e.g., Farmers). Relationships (e.g., HAS BANK, HAS AGENCY, PRODUCES) between these entities are also visualized.

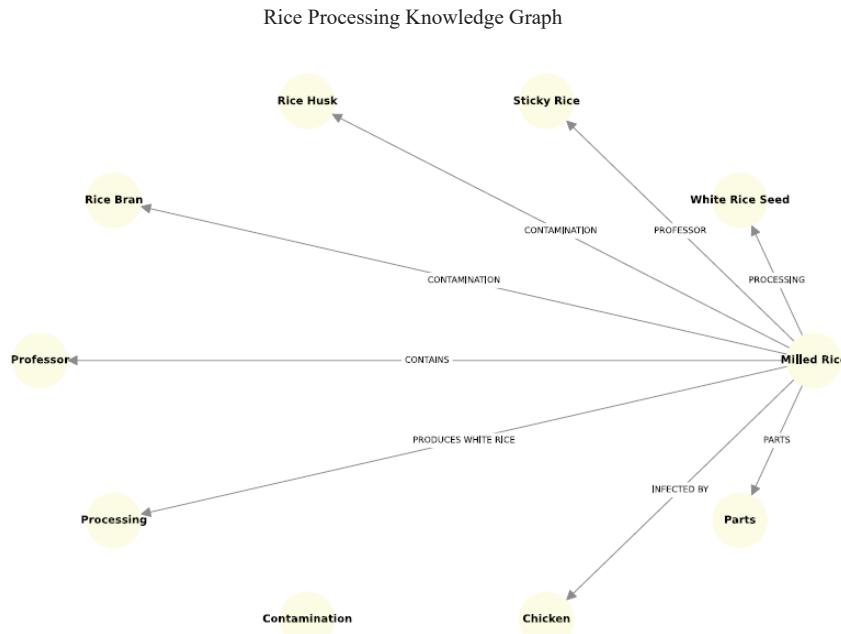


Fig. 4. Rice Processing Knowledge Graph. It visualizes the knowledge graph focused on rice processing stages and related entities. Entities include Rice Husk, Milled Rice, Sticky Rice, Rice Bran, and White Rice Seed, as well as concepts like Processing and Contamination. The Relationships shown are CONTAINS, PRODUCES WHITE RICE, INFECTED BY, and CONTAMINATION.

B. How the Knowledge Graph is Set Up

The GraphRAG system uses a knowledge graph, which we can think of as $G = (V, E)$. Here, V is like a list of nodes (or things), such as different kinds of rice, types of soil, and ways of growing rice. And E is like a list of edges (or relations) that show how these things are connected. These lines have labels to say what kind of connection it is. For example, a line $e = (v_i, \text{PART_OF}, v_j)$ could mean that a type of rice v_i is part of a certain group or has a special feature linked to v_j [17].

This setup lets GraphRAG use Neo4j's labeled lines to create and follow connections easily. This helps it understand complex information about Thai rice. We can also think of the knowledge graph using a table called an adjacency matrix, where each element is:

$$A_{ij} = \begin{cases} 1 & \text{if there is the line from thing } v_i \text{ to thing } v_j \\ 0 & \text{otherwise} \end{cases}$$

This helps GraphRAG find connected pieces of information that are needed to answer questions, especially when questions are about specific topics where connections are very important for giving good answers [17].

C. Using Cipher Language for Math

To make GraphRAG better at understanding connections in the knowledge graph, we use Neo4j's Cypher language to do some math on the graph. Cypher lets GraphRAG find complex connections,

gather information together, and do calculations on the dots and lines in the graph [14].

For example, to find the **total strength** of connections between linked dots (showing how strong or common connections are), we use this Cypher command:

```
MATCH (n) - [r: RELATIONSHIP_TYPE] -> (m)
RETURN n.name, SUM (r.weight) AS totalWeight
```

This command adds up the strengths of all connections between dots (n) and (m) for a certain type of connection. This helps GraphRAG measure and understand how important these connections are when it makes answers [17].

Also, to find the average distance between dots in the knowledge graph (which shows how close things are, like how many steps between a rice type and its features), we use this Cypher command:

```
MATCH p = (a)-[*] -> (b)
RETURN a.name, b.name, AVG (length (p))
AS avgPathLength
```

This command calculates the average distance across different paths between dots. This helps GraphRAG figure out how close and relevant things are when it is answering complex questions that need context [17].

The way we did this study is focused on making and testing the GraphRAG system. GraphRAG is a tool that uses knowledge to help answer questions about Thai rice using a method called Retrieval-Augmented Generation (RAG). The system uses several advanced tools, including Neo4j to manage the knowledge graph, LangChain to get text from PDF files, and the Llama 3 8b Large Language Model

(LLM) to create natural-sounding language. The parts below explain the main pieces of this system and how we tested how well it works.

D. Getting the Data

To create a strong information base, GraphRAG used data from trusted sources for Thai farming information. We got important details from PDF files. These files gave us the key facts needed to answer questions about Thai rice:

- Thai Rice Department: Gave basic info about rice types, how to grow rice, and rules from the government that affect the Thai rice business.
- Thai Agriculture Department: Gave data on taking care of soil, controlling bugs and plant diseases, and good farming methods that help the environment, which are all important for growing rice.
- Kasetsart University, Agriculture Faculty: Shared research from scientists on making better types of rice, improving rice plants through genetics, and getting the best rice harvests.
- Rice Disease Info Files: Gave specific help on finding, handling, and stopping rice diseases.
- Other PDF Files: Had info on rules for selling rice to other countries, how the rice market is changing, and how to deal with climate change in Thai rice farming.

We used LangChain to take the text out of these files and then organized it into a knowledge graph using Neo4j. This helps GraphRAG give clear and correct answers about farming methods, rice plant health, and money matters in Thai rice growing.

E. Examples for Testing

Besides the numbers, we also asked people what they thought. People liked GraphRAG because it gave short, focused answers, especially for hard questions that needed to know how different things are connected. However, for some simpler questions where people expected more explanation, they liked the normal RAG systems better because they gave more details. Knowing this helps us make GraphRAG better in the future. We need to find the right balance between giving quick answers and giving enough information.

Example 1

Question: พื้นที่ปลูกข้าวที่ใหญ่ที่สุดอยู่ที่ไหน? (Where is the largest rice-growing area?)

RAG Response: Thai rice is grown throughout the country, but the largest rice-growing areas are in the Central and Northeastern regions.

GraphRAG Response: India.

Feedback: GraphRAG is better—The user preferred the concise answer from GraphRAG, despite it being factually incorrect. However, the user's preference was likely due to the perceived directness of the response.

Example 2

Question: แนวโน้มการผลิตข้าวในประเทศไทย? (Rice production trends in Thailand?)

RAG Response: Between 1985 and 2020, rice production per hectare of Thai farmers increased by an average of 0.8.

GraphRAG Response: Thailand has a tendency to increase rice production in 2022 with a rice production volume of approximately 15.5 million tons.

Feedback: RAG is better—The RAG system provided a more detailed and contextually appropriate definition, whereas GraphRAG's response was too brief and vague.

Example 3

Question: ข้าวมีกี่สายพันธุ์ ต่างกันอย่างไร? (How many varieties of rice are there? How are they different?)

RAG Response: There are about 40-50 types of Thai rice. Each type of rice has different colors, flavors, and characteristics.

GraphRAG Response: Rice has many species, divided into 2 large groups: *Oryza sativa* group, which is rice used for direct consumption, and *Oryza glaberrima* group, which is rice used for making glutinous rice. Rice used for direct consumption is divided into 3 species: Sticky rice, which is sticky and thick; Aromatic rice, which is fragrant and has a sweet taste. Regular rice is a mixture of Sticky rice and Aromatic rice. Each species has different characteristics and properties. Glutinous rice is used to make Sticky rice. It is useful for making Sticky rice.

Feedback: GraphRAG is better—The GraphRAG System provides more descriptive and specific species names than the RAG, which provides general information.

Example 4

Question: ข้าวหอมมะลิแดงที่ปลูกในภาคใต้และเจอน้ำท่วมขัง มีวิธีการป้องกันและดูแลข้าวให้ปลอดภัยและไม่เสียหายหรือไม่? (Red jasmine rice grown in the South that is flooded; is there a way to protect and care for the rice so that it is safe and not damaged?)

RAG Response: Yes, there are several ways to protect and care for this red sticky rice, such as using good drainage systems and controlling water levels to the right level.

GraphRAG Response: Red jasmine rice planted in the South that is flooded can be protected and cared for safely and without damage by the following methods: 1) Soil preparation before planting, prepare the soil for rice planting by cleaning the soil from weeds and other materials and adding fertilizer to the soil to increase soil moisture and temperature. 2) Planting rice in a position higher than the water level: Plant rice in a position higher than the water level to prevent water from flooding. 3) Using a drainage system: Use an appropriate drainage system to drain

water that is stuck on the ground. 4) Caring for rice after planting: Care for rice after planting by cleaning the soil of weeds and other materials and adding fertilizer to the soil to increase soil moisture and temperature. 5) Monitoring for diseases and insects: Monitor for diseases and insects that may occur with rice and use appropriate methods to control diseases and insects. Overall, preventing and caring for red

jasmine rice planted in the South that is flooded requires caution and patience in caring for the rice consistently.

Feedback: GraphRAG is better – The user found GraphRAG's response provides more useful tips and methods in a step-by-step format, which makes it more understandable to the user.

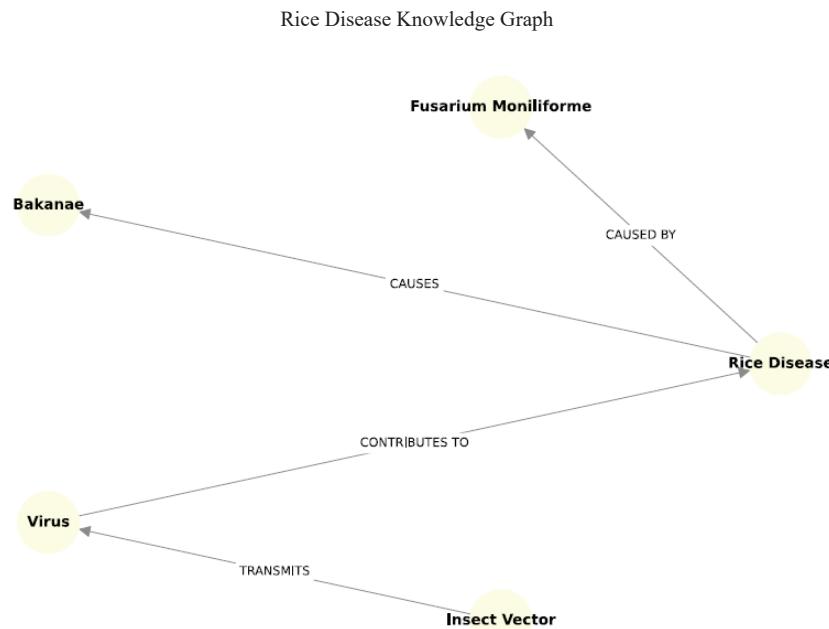


Fig. 5. Rice Disease Knowledge Graph. It visualizes the relationships between different factors related to rice diseases. Entities include types of diseases (e.g., Bakanae, Rice Disease), pathogens (Fusarium Moniliforme, Virus), and vectors (Insect Vector). Relationships depicted are CAUSES, CONTRIBUTES TO, TRANSMITS, and CAUSED BY.

F. Limitation and Future Works

Even though GraphRAG is better than older systems in some ways, it's not perfect. Sometimes, it has trouble with questions that need really deep science knowledge or questions that are unclear and can be understood in different ways. To make it better, we will work on:

Making the Information Base Bigger: We will add more information to the knowledge graph so it knows more things. **Getting Better at Understanding Tricky Questions:** We will teach the system to be better at figuring out questions that are not clear or have multiple meanings. **Improving How Answers Are Made:** We will make the system better at creating answers so it can handle many different kinds of questions. Also, we want to use GraphRAG for other types of farming, not just rice. This will help us see if it can work well in different areas [9].

Market Trends and Economics: Questions in this area addressed economic aspects of rice farming, such as:

- What is the current market price for rice?
- What is the demand for organic rice?

Export and Trade Regulations: This category covered the legal and logistical aspects of rice export, including:

- What are the regulations for exporting rice to different countries?
- What is the certification process for exporting organic rice?

IV. SYSTEM ARCHITECTURE

The GraphRAG system was developed using a modular architecture to ensure scalability and flexibility in query handling. To effectively compare GraphRAG with a traditional approach, we implemented both GraphRAG and a baseline RAG system, each leveraging distinct data storage and retrieval mechanisms. The system architectures are visualized as shown in Fig. 6 and detailed in the following subsections.

1) *LangChain for Text Extraction (Both Systems):* Both the GraphRAG and baseline RAG systems utilize LangChain for extracting relevant textual data from unstructured PDF documents related to Thai rice. Specifically, we employed the *PymuPDF*

module to process PDF documents and the *Langchain core.documents* for chunking text. This ensures consistent and effective text preparation across both systems, focusing the comparative evaluation on the impact of knowledge graph integration [11].

2) *Neo4j Knowledge Graph Database (GraphRAG)*: The GraphRAG system's core innovation lies in its Neo4j knowledge graph. This graph database is optimized for managing and querying large-scale, interconnected data. Neo4j structures the extracted Thai rice data into entities and relationships, as described previously, enabling GraphRAG to perform graph-based retrieval for contextually rich information. Cypher query language is used for efficient and expressive interaction with the Neo4j database [17].

3) *Pinecone Vector Database (Baseline RAG)*: In contrast to GraphRAG, the baseline RAG system utilizes Pinecone as its data storage and retrieval mechanism. Pinecone, a vector database, stores vectorized embeddings of the extracted text chunks. For the baseline RAG, we employed sentence-transformers/all-mnlp-base-v2 from Hugging Face embeddings to generate these embeddings. Retrieval in the baseline RAG system is performed using a similarity search over these vector embeddings, a common approach in traditional RAG pipelines.

4) *Llama 3 8b Language Model via Groq API (Both Systems)*: For response generation in both GraphRAG and the baseline RAG systems, we

employed the Meta Llama 3 8b Instruct model, accessed through the Groq API. Utilizing the Groq API provides low-latency inference, crucial for real-time query responsiveness and for facilitating isolating the impact of the knowledge graph versus vector database retrieval on response quality. Specific prompting strategies were designed to guide the LLM to generate informative and domain-specific answers based on the retrieved context [18].

5) *Gradio Interface for preference test and System Evaluation*: To facilitate user evaluation and performance analysis, we developed a Gradio interface. This interface enabled a side-by-side preference test of GraphRAG and the baseline RAG system, allowing participants to directly compare the responses for the same queries. Furthermore, the Gradio interface was instrumented to measure key performance indicators, such as model response time (latency) and memory usage for both systems. These metrics were collected to provide quantitative data on system efficiency alongside qualitative user feedback. The process starts with taking Thai text from PDF files about rice using LangChain. Then, the text is broken into smaller pieces and put into a knowledge graph using Neo4j. When someone asks a question, the system finds the right information from the knowledge graph and sends it to the Llama 3 8b model through the Groq API. Then, the model creates an answer that makes sense with the question.

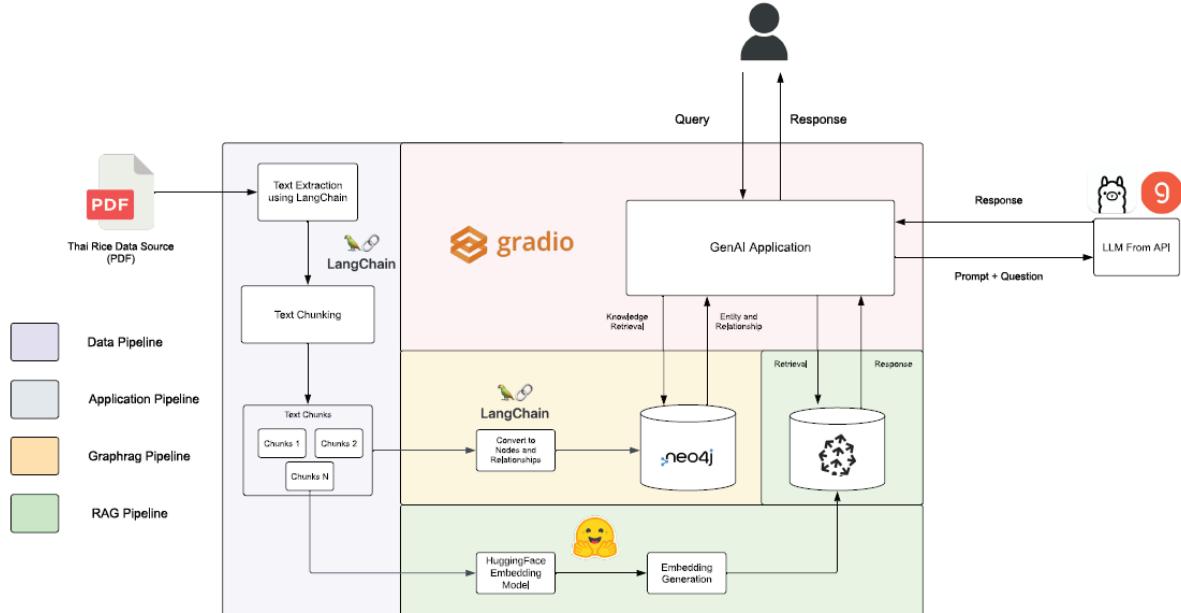


Fig. 6. System Architecture for Thai Rice Assistant. The system is divided into four main pipelines: Data Pipeline, Application Pipeline, GraphRAG Pipeline and RAG Pipeline. Each pipeline is color-coded to highlight its specific role in data processing and retrieval.

V. RESULTS

We performed preference tests on our two different systems using 102 questions. We compared GraphRAG with a basic RAG system. We got feedback from three groups of people: Thai farming experts, researchers

studying for their Master's degree at Tuxsa, and AI students from the university. These people helped us see how well GraphRAG worked in real situations.

The answer choices of this test are A was Better, B was Better, both acceptable, or both unacceptable. The results of what they thought are in Table I. This is

a blind evaluation. Users will not know the identities of systems A and B. To see how well GraphRAG worked, we tested it against a normal system for getting information called Retrieval-Augmented Generation (RAG). Normal RAG just finds information from messy sources without much understanding. But GraphRAG uses organized links between things like rice types, farming methods, and weather conditions. Because of these connections, they give answers that are more correct and understand the situation better. The results of this experiment are shown in Table I. To evaluate the effectiveness of the proposed GraphRAG system against the baseline RAG, we employed both quantitative performance metrics and qualitative user-based assessment. Performance was measured in terms of average response time and system memory usage during query processing. For qualitative evaluation, we conducted a user preference test where participants rated the relevance, completeness, and overall helpfulness of responses generated by both systems for a predefined set of queries representing common information needs of Thai rice farmers. **While standard metrics like precision and recall are challenging to compute without a large-scale, manually annotated ground-truth dataset for this specific domain, the user preference scores serve as a practical measure of the system's perceived correctness and utility in addressing real-world queries.** Furthermore, the inherent structure of the knowledge graph used in GraphRAG potentially allows for greater explainability compared to standard RAG, as the system can trace information back to specific entities and relationships within the graph [19].

TABLE I
PREFERENCE TEST RESULTS: FEEDBACK ON SYSTEM
RESPONSE PREFERENCE

Result	Percentage
RAG	35.3%
GRAPHRAG	52.9%
BOTH ACCEPTABLE	6.9%
BOTH UNACCEPTABLE	4.9%

Here are some examples to show why people liked GraphRAG better – because its answers were more useful and right on point:

Question 1: การควบคุมพืชเมืองกีการใดบ้างโดยไม่ต้องใช้สารเคมี? (What are some methods of controlling plants without using chemicals?)

RAG: Good rice seeds will produce good results. If the seeds are not good, have diseases, or are mixed with a lot of weeds, the results will be less.

GraphRAG: Choose good rice varieties with characteristics such as high yield, weather resistance, soil resistance, water resistance, and beauty. The normal RAG answer was not clear. It talked about

good seeds but didn't say anything about controlling plants without using chemicals. But the GraphRAG answer was much better. It gave real steps that farmers could take, so it fit the question better. People liked this answer because it was easy to understand and helpful.

Question 2: ข้าวพันธุ์ใดให้ผลผลิตสูงสุด? (Which rice varieties produce the highest yields?)

RAG: Good rice varieties will produce good results because good-quality rice seeds will produce good results.

GraphRAG: Rice variety 105

The normal RAG answer kept saying the same thing over and over and wasn't very specific. But GraphRAG was straight to the point. It said exactly "Rice variety 105", which was a clear and right answer. People liked GraphRAG because it gave direct and useful answers.

User Preference and Feedback

Overall, people thought the GraphRAG answers were better because they were more specific and helpful. Because GraphRAG has a knowledge graph, it can use organized information and connections about Thai rice. This means its answers give you real information that you can use. This shows that GraphRAG is better than normal RAG systems for helping with questions about Thai rice.

TABLE II
AVERAGE RESPONSE TIME AND MEMORY USAGE PER
QUESTION

System	Response Time (s)	Memory Usage (KB)
RAG	1.43	457.42
GraphRAG	0.41	213.09

Table II shows the comparison of average response time and memory usage per question between the traditional RAG system and GraphRAG. GraphRAG demonstrates a significant improvement in efficiency, with a response time approximately three times faster than RAG (0.41 seconds compared to 1.43 seconds) and a reduction in memory usage by more than half (213.09 KB compared to 457.42 KB).

The results, particularly the strong user preference for GraphRAG responses and its ability to handle complex queries more effectively than the baseline RAG system, align with the growing body of research highlighting the benefits of incorporating structured knowledge into generative AI systems [20]. Our findings support the hypothesis that leveraging explicit relationship information stored within a knowledge graph allows GraphRAG to provide more contextually relevant and accurate answers, confirming the advantages observed in GraphRAG applications in other domains and demonstrating its effectiveness even within the specific constraints of the low-resource Thai rice information ecosystem.

VI. CONCLUSION

This research demonstrated the successful application of GraphRAG, a knowledge graph-enhanced RAG approach, to improve query assistance for Thai rice cultivation. The system was developed using a corpus of 6 key documents related to Thai rice farming practices, resulting in a knowledge graph comprising approximately 200 chunks, which consist of 3000 maximum words per chunk. Our comparative analysis showed that GraphRAG significantly outperformed a traditional RAG baseline in user preference tests, particularly for complex queries, while maintaining reasonable performance efficiency. This highlights the value of integrating structured knowledge graphs to enhance contextual understanding and information retrieval accuracy, especially in specialized domains like agriculture.

The tests we did showed that GraphRAG is better than basic systems, especially when dealing with hard questions. These hard questions need a good understanding of how farming works, how rice is sold to other countries, and how different countries compete. By using information from places where there isn't much data, GraphRAG has shown it can be good at finding information even when it's hard to get data.

Future work will focus on several key areas. Firstly, we plan to **expand the knowledge graph** by incorporating a wider range of data sources, including real-time weather data and market prices, to provide more comprehensive and timely information. Secondly, we aim to improve the system's natural language understanding capabilities to better handle ambiguous or colloquial queries from farmers. Thirdly, enhancing the generative component to produce more nuanced and actionable advice is a priority. Finally, exploring the adaptability of this GraphRAG framework to other agricultural domains within Thailand represents a promising avenue for future research.

ACKNOWLEDGMENT

The authors appreciate the support of the Tuxsa Master's Degree Research Group at Thammasat University for helping with data collection and to all participants who contributed critical feedback to this investigation.

REFERENCES

- [1] D. S. Birla, K. Malik, M. Sainger, D. Chaudhary, R. Jaiwal, and P. K. Jaiwal, "Progress and challenges in improving the nutritional quality of rice (*Oryza Sativa L.*)," *Crit. Rev. Food Sci. Nutr.*, vol. 57, no. 11, pp. 2455-2481, Jul. 2017.
- [2] R. Wassmann et al., "Climate change affecting rice production: The physiological and agronomic basis for possible adaptation strategies," *Adv. Agron.*, vol. 101, pp. 59-122, Oct. 2009.
- [3] K. Kiratiratanapruk et al., "Development of paddy rice seed classification process using machine learning techniques for automatic grading machine," *J. Sensors*, vol. 2020, no. 1, p. 7041310, Jul. 2020.
- [4] P. P. Bhat, R. Prasad, K. Anil, and A. Jadhav, "The role of information and communication technology in enhancing the effectiveness of agricultural extension programs worldwide: A review," *J. Sci. Res. Rep.*, vol. 30, no. 7, pp. 963-976, Jul. 2024.
- [5] W. Nwankwo, C. P. Nwankwo, and A. Wilfred, "Leveraging on Artificial Intelligence to accelerate sustainable bioeconomy," *IUP J. Knowl. Manag.*, vol. 20, no. 2, pp. 35-59, Apr. 2022.
- [6] C. Ling et al., "Domain specialization as the key to make large language models disruptive: A comprehensive survey," *Journal of Economics and Management*, vol. 46, 509-583, Dec. 2023.
- [7] Y. Huang and J. Huang, "A survey on retrieval-augmented text generation for large language models." arXiv, 2024. [online]. Available: <https://arxiv.org/abs/2404.10981> [Accessed Apr. 2024].
- [8] S. Rezayi et al., "Exploring new frontiers in agricultural NLP: Investigating the potential of large language models for food applications," *IEEE Trans. Big Data*, 2024.
- [9] Q. Zhang et al., "A survey of graph retrieval-augmented generation for customized large language models." arXiv, 2025. [online]. Available: <https://arxiv.org/abs/2501.13958> [Accessed Jan. 2025].
- [10] B. Peng et al., "Graph retrieval-augmented generation: A survey." arXiv, 2024. [online]. Available: <https://arxiv.org/abs/2408.08921> [Accessed Aug. 2024].
- [11] S. Ghane, R. Sawant, G. Supe, and C. Pichad, "LangchainIQ: Intelligent content and query processing," *Int. J. Manag. Technol. Soc. (IJMITS)*, vol. 9, no. 3, pp. 34-43, 2024.
- [12] ArangoDB, "ArangoDB's GraphRAG Transforms Healthcare Data Management," *ArangoDB Blog*, 2025. [online]. Available: <https://arangodb.com/2025/04/arangodb-graphrag-transforms-healthcare-data-management/> [Accessed Apr. 6, 2025].
- [13] T. Balarabe, "GraphRAG vs Traditional RAG: Knowledge Graphs for Accurate, Enhanced RAG Applications," Medium. com, 2025. [online]. Available: <https://medium.com/@tahirbalarabe2/graphrag-vs-traditional-rag-knowledge-graphs-for-accurate-enhanced-rag-applications-2cc4f6f9f4b4> [Accessed: Apr. 6, 2025].
- [14] A. N. Arifin, "Are Re-Ranking in retrieval-augmented generation methods impactful for small agriculture qa datasets? a small experiment," *BIO Web of Conf.*, vol. 167, pp. 1-6, Mar. 2025, <https://doi.org/10.1051/bioconf/202516701001>
- [15] C. Pechsiri and R. Piriaykul, "Developing the UCKG-WHY-QA system," in *Proc. 7th Int. Conf. Comput. Converg. Technol. (ICCCT)*, IEEE, 2012, pp. 679-683.
- [16] Y. Xie, L. Jia, and J. Dai, "Construction of a traditional Chinese medicine Dao Yin science knowledge graph based on Neo4j," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, IEEE, 2023, pp. 4662-4666.
- [17] C. Johnpaul and T. Mathew, "A Cypher query based NoSQL data mining on protein datasets using Neo4j graph database," in *Proc. 4th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, IEEE, 2017, pp. 1-6.
- [18] R. Sapkota et al., "Multi-modal LLMs in agriculture: A comprehensive review," *TechRxiv*. [Online]. Available: <https://www.techrxiv.org/users/795986/articles/1224079-multi-modal-lmms-in-agriculture-a-comprehensive-review> [Accessed Sep. 2024].
- [19] Lettria, "10 rules for optimizing your GraphRAG strategies," *Lettria Blog*, 2025. [Online]. Available: <https://www.lettria.com/blogpost/10-rules-for-optimizing-your-graphrag-strategies> [Accessed Apr. 6, 2025].
- [20] AWS, "Improving Retrieval Augmented Generation accuracy with GraphRAG," *AWSMachineLearningBlog*, 2025. [Online]. Available: <https://aws.amazon.com/blogs/machine-learning/improving-retrieval-augmented-generation-accuracy-with-graphrag/> [Accessed Apr. 6, 2025].



Language Models.

Gampanut Soontontam is currently a student in the Bachelor of Computer Engineering and Artificial Intelligence at Panyapiwat Institute of Management. His research interests are in the area of Machine Learning and Large



Kanabadee Srisomboon is an Associate Professor of the Department of Electrical and Computer Engineering at King Mongkut's University of Technology North Bangkok, Bangkok, Thailand. He received his B.S. degree from Civil Aviation Training Center in 2011. The M.Eng. and Ph.D. degrees from King Mongkut's University of Technology North Bangkok were received in 2013 and 2016, respectively. His research interests include image processing, smart grid communication, Cognitive Radio Communication, and Artificial Intelligence technology.



Tinnaphob Dindam is an Assistant Professor in Electrical Engineering and Lecturer in Computer Engineering and Artificial Intelligence at Panyapiwat Institute of Management, Nonthaburi, Thailand.

He graduated with a B.S.Tech. Ed., M.Eng., and Ph.D. in Electrical Engineering from King Mongkut's University of Technology North Bangkok. His research focuses on wireless networks, cognitive radio, and Machine Learning.



Adisorn Kheaksong is an Assistant Professor in Computer Engineering and Lecturer in Computer Engineering and Artificial Intelligence at Panyapiwat Institute of Management, Nonthaburi, Thailand.

He received his B.Ind. Eng. in Electrical Engineering from Saint John University, M.Sc. in Computer Science from Rangsit University, and Ph.D. in Electrical Engineering from King Mongkut's University of Technology North Bangkok. His research focuses on wireless networks, cloud computing, and emerging technologies in artificial intelligence.



Parinya Sanguansat is an Associate Professor in Electrical Engineering and head of Computer Engineering and Artificial Intelligence at Panyapiwat Institute of Management, Nonthaburi, Thailand.

He graduated B.Eng., M.Eng. and Ph.D. in Electrical Engineering from Chulalongkorn University. His research areas include Machine Learning, Computer Vision, and Artificial Intelligence.