International Review of Management and Marketing

ISSN: 2146-4405

available at http: www.econjournals.com

International Review of Management and Marketing, 2025, 15(1), 61-70.

Integrating Management and Marketing Strategies in Organic Farming: A Topic Modeling Analysis of Sustainable Development and Entrepreneurship

Sumana Chiangnangam, Paiboon Manorom, Wirapong Chansanam*

Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand. *Email: wirach@kku.ac.th

Received: 23 August 2024 **Accepted:** 14 November 2024 **DOI:** https://doi.org/10.32479/irmm.17559

ABSTRACT

Integrating management and marketing with organic farming practices is crucial for the sustainable development of this growing sector. As organic farming products are popular among health-conscious consumers, agricultural entrepreneurs increasingly use innovative business strategies to meet the rising demand. The successful alignment of business and economic principles with organic agriculture enhances value creation and promotes long-term sustainability. Entrepreneurs must focus on strategic networking across production, processing, and marketing to capitalize on these opportunities fully. Despite its growth, the organic agriculture sector still faces challenges that require comprehensive research to develop effective strategies for advancing the industry. This study addresses the intersection of organic agriculture, entrepreneurship, and business strategy by utilizing topic modeling techniques to analyze bibliographic data from 4327 articles published between 1946 and 2023 in the Scopus database. Using Latent Dirichlet Allocation (LDA) topic modeling, five key themes were identified: (1) Soil for organic agriculture, (2) Environment and organic agriculture, (3) Agriculture business, (4) Organic production, and (5) The use of organic substances. With a connection value of 0.419, the topic modeling effectively grouped relevant themes, offering valuable insights into the role of management and marketing in organic farming. These findings provide a solid foundation for future research and innovation, enabling entrepreneurs, researchers, and farmers to align their practices with evolving market trends better, thus fostering the continued growth of organic agriculture.

Keywords: Organic Agriculture, Entrepreneurship, Sustainable Development, Topic Modeling, Agricultural Marketing
JEL Classifications: M0, M1, Q1

1. INTRODUCTION

Agriculture in Thailand is evolving by technology continuously and has adapted innovation to transform into the era of Agriculture 4.0, which emphasizes high-quality agriculture to increase production productivity in quantity and value of agricultural products (Office of the Small and Medium Enterprises, 2020). Furthermore, the government sector also expects farmers to practice sustainable agriculture in three dimensions, i.e., the economy, the society, and the environment, to create growth based on an environmentally friendly (Research Institute of Organic Agriculture, 2021) through promoting agricultural products and processed agricultural products to create high value, which is environmentally friendly and build an image of Thailand to be tourism destination that emphasizes value and sustainability (National Organic Agriculture Development Committee, 2017).

Organic farming is considered a sustainable agricultural system, which gained a number of attention from farmers and consumers and resulted in the organic agricultural products have increased along with the project for promotion and support for farmer groups to enter organic farming and strengthen farmer groups by connecting networks in production, processing, and marketing. Particularly, in the National Economic and Social Development

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Plan No. 13 (2023-2027), was mentioned the objectives related to the development of organic agriculture as follows: (1) High-value economy with friendly environment and to be a leading country in agricultural products and agricultural processing that create high value and build Thailand to be a tourism destination that emphasizes value and sustainability, (2) Society of opportunity and equality for strong SMEs with high potential and high competitiveness, (3) Sustainable lifestyles for a circular economy and a low-carbon society to reduce risks and impacts from natural disasters and, (4) Factors Driving Development for a highly capable workforce for future development and needs (Announcement on the National Economic and Social Development Plan No. 13, 2022). In addition, the government has urgent policies to lay the foundations of the country's economic system for the future of technology development along with developing the skills of small and medium-sized entrepreneurs (SMEs) and communities as well as urgent policies for developing public services systems to develop the government's data collection and disclosure system to develop a central organic agricultural database for the entire organic agricultural supply chain and support research and development to create systematic knowledge and innovation in a concrete way (Agricultural Research Development Agency, 2018; Office the Small and Medium Enterprises, 2020).

Before getting to know the entrepreneur, there are numerous information that might need to be understood in depth to be used as a guideline to drive the growth of entrepreneurial groups, such as production system information, production factors, market intelligence, technology, innovation information, and so on. Due to the collection of organic agricultural information should be categorized in the same way and rely on the updated resources of information, knowledge, and skills (Blake and Wijetilaka, 2015), with researchers in various fields conducting research on organic farming and entrepreneurship and there were a number of studies have listed in the database. However, finding useful information or additional information about organic agriculture and entrepreneurs seems to be difficult to retrieve information into the public sector (Joo et al., 2016). Furthermore, analyzing large amounts of data also requires tools to help in aspect processing information, e.g., natural language processing, which analyzes huge amounts of text from social media data, including topic modeling techniques (Jelodar et al., 2019).

Topic modeling is a model of data distribution for categorization and contains ideas from various documents, resulting in a huge amount of data and a collection of topics. Each topic has a probability of words occurring in that topic. Thus, the topic modeling is the creation of a data distribution model which was used in grouping data based on the idea that a document is a combination of huge resources. Consequently, the topics might have a probability distribution of many words occurring in each topic based on the concept of the Latent Dirichlet Allocation: LDA, which was created with the idea that a document would consist of topics together randomly in the form of word groups (Diego et al., 2020). In the aspect of searching, LDA is used to find topics or word groups that need to be extracted from the document by calculating the probability value (probabilistic) from words appearing in documents as a latent topic, which cannot be clearly observed. Then, the LDA program will analyze the probability of each word in the latent topic and estimate the proportion of hidden topics in a document. The LDA is the most commonly used which comprises of a flexible method and changeable for creating topic models to separate important points of the message. To create a topic model, the question can be answered: Which topics are discussed most frequently? Therefore, topic modeling will provide more insight into relevant topics within that topic by creating content relevance through selecting semantically relevant keywords (Donthu et al., 2021). Meanwhile, the content is then analyzed to determine whether it is relevant or not, which is widely used in various fields of study such as Linguistic science, Political science, Medical and Biomedical, Geographical and locations, etc. (Jelodar et al., 2019). Hence, this study is an analysis of research on organic agriculture and entrepreneurs in the relevant context and citation, which relevant topics are discussed the most in the field of study.

2. RESEARCH METHODOLOGY

This study has employed a quantitative analysis of bibliographic research articles through the text analysis method from research abstracts with the aim to understand the research trends of organic agriculture, entrepreneurship, and knowledge pedagogy. Nonetheless, the application of this technique uses the concept of text analysis with a topic model to analyze topic model creation. Then, the researcher considered the prominent words of each issue and grouped them together for the important issues that correspond to the others. The research methodology was comprised of four stages, as were as follows:

2.1. Data Collection

In this study, the researcher has searched for documents related to organic agriculture and entrepreneurs in the Scopus database by specifying keywords used in the search. Then, analyzed the main issues regarding organic agriculture and agricultural entrepreneurs in the heading or title of the various sources, such as documents, books, articles, and related research, to identify the keywords. After that, proceed with the obtained words in the search using advanced search techniques with AND to connect the words to get the results in which both words appear. Besides, the OR was used to connect words to get the results in which either word appears (Wangsunthornchai, 1998). Therefore, the research technique would be "(organic AND agriculture) OR (organic AND farming) OR (organic AND business) OR (organic AND startup) OR (organic AND entrepreneur)". The method has selected relevant documents to organic agriculture, entrepreneurs, and businesses. Thus, the appeared document would be indexed in the fields of Social Sciences, Business Management and Accounting, Economics/Econometrics and Finance, Multidisciplinary, Arts and Humanities and has been published from 1946 to 2023 in English only. The data in this study was downloaded on April 11, 2023 and found that 6730 results were obtained. Then, the researcher checked for duplication, anonymous author, and abstract absence to be excluded from this study. The final results were about 4,327 in total and were saved in the file as a CSV format for data analysis.

Table 1: The topic modelling from the representative topic

Topic No.	The representative topic (bigrams)	Topic Name	Description
0	Organic farming, organic matter, land use, soil organic, farming system, organic carbon, result show, organic conventional, soil quality, soil fertility, organic fertilizer, cropping system, soil property, farming practice, improve soil, management practice, production system, crop production, crop yield, conventional farming	Soil for Organic Agriculture	The study organic agriculture requires studying soil in areas related to organic agriculture, such as land use, soil quality, and soil properties. soil improvement adding nutrients to the soil, soil management, etc.
1	Environmental impact, organic farming, production system, result show, farming system, land use, organic waste, case study, energy consumption, management system, organic production, waste management, environmental performance, dairy farm, energy efficiency, energy use, milk production, conventional organic, use organic, organic conventional	Environment and Organics	The origins of organic farming were intended to preserve the environment to reduce the impact on the environment, namely agriculture using chemicals. This resulted more than chemical residues in the soil mixed in water and air continuously affecting the health of consumers both directly and indirectly. A study of the article shows that this group of liquor information has emerged to have more knowledge and understanding of organic farming.
2	Organic farming, case study, organic agriculture, organic food, food system, organic growth, sustainable development, business model, organic farmer, small business, organic farm, purpose paper, food production, organic production, development organic, organic product, organic sector, result show, business environment, social movement	Organic agriculture business	The organic farming relates to the business sector. Currently, there are many forms of organic farming business according to the results of the topic analysis in the article, related to management, production, development, inspection, and strategy were business systems that classify small businesses, kitchen business, new business, including presentation of business results.
3	Organic farming, organic food, organic agriculture, organic product, organic production, organic farm, organic farmer, farming system, result show, organic conventional, food system, case study, farming practice, food product, land use, food production, production system, conventional organic, conventional farming, sustainable agriculture	Organic production	The organic farming has consisted of many factors to get quality organic products and maintains organic condition about organic agriculture, organic products, organic producer, Safe-food production, food systems, food products, including organic certification and agricultural policy.
4	Organic farming, organic solvent, organic acid, result show, method apply, organic conventional extraction method, first time, result indicate, present study, volatile organic, study aim, method use, result obtain, organic residue, optimum condition, high quality, ground state successfully apply	Organic use	Having the organic farming knowledge about agriculture might not enough but may need to have knowledge about organic substance to be able to use various compounds correctly, which does not violate the principles of organic farming. The article therefore studies and publishes content about organic substances usage, methods for extracting organic compounds, residues, on-organic and organic to provide those interested in studying organic agriculture with guidelines for further study.

Figure 1: Published research article articles from 1946 to 2023

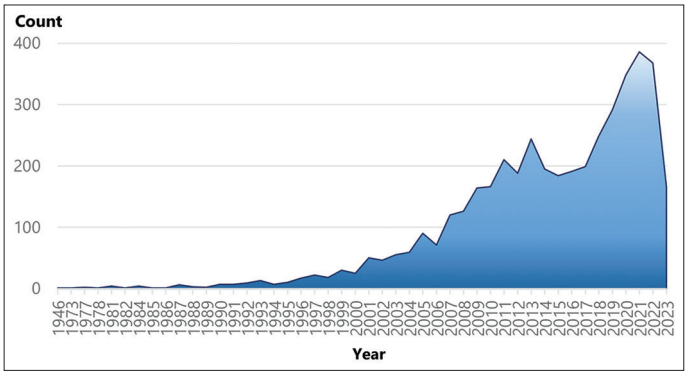
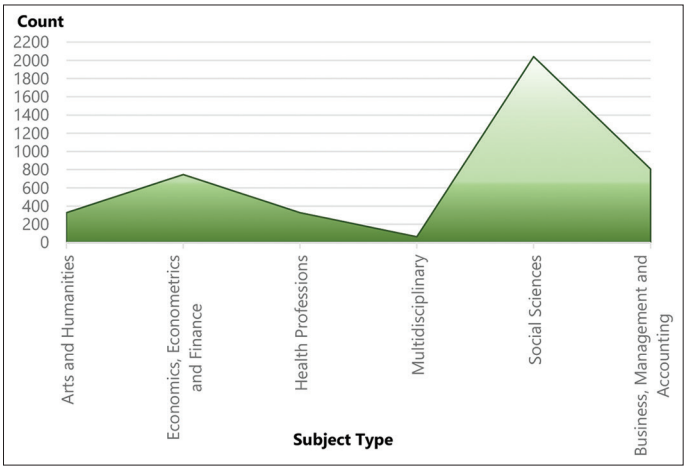


Figure 2: Article articles in each topic



2.2. Data Preparation and Data Cleaning with Natural Language Processing (NLP)

For the data preparation and cleaning, the data was imported in the English language through the Natural Language Processing (NLP) at the pre-processing step to clean the data, which accelerates computers to understand human language. The NLP would assist in categorizing, summarizing, and creating text, which converts the data with reliable tools as open source and available to run on Python. The text at the pre-processing proceeded in PyCaret (Ali, 2020). These steps involved deleting commonly used but insignificant words, which would help to delete the marks, punctuation, numbers, and unnecessary words to extract representative keywords from articles.

Figure 3: Data from imported file. CSV format

```
display(dataset)
```

	Authors	Title	Year	Cited by	Abstract	Sbj. A.
0	A.V V.B.; Baresel J.P.; Weedon O.; Finckh M.R.	Effects of ten years organic and conventional ...	2019	13	Early vigour traits of wheat composite cross p...	0
1	Aaijaz N.; Bin Ibrahim M.D.; Bin Ahmed G.	Green consumers: A growing market for SME'S an...	2010	1	This study attempts to gain knowledge about th...	3
2	Aarnink A.J.A.; Hol J.M.G.; Beurskens A.G.C.	Ammonia emission and nutrient load in outdoor ...	2006	16	Ammonia emission and nutrient load in outdoor ...	4
3	Aarts H.F.M.; Conijn J.G.; Corré W.J.	Nitrogen fluxes in the plant component of the ...	2001	14	Sandy areas in th...	4
4	Abbas A.; Sajid M.B.; Sajid J.; Ahmed N.	Forecasting environmental and social benefits ...	2023	1	Embodied carbon of new buildings can be effect...	4
...	...	...	...	...	...	...
4322	Zrakić M.; Jež Rogelj M.; Grgić I.	Organic agricultural production on family farm...	2017	7	The Croatian organic products market is not de...	4
4323	Zuba-Ciszewska M.; Kowalska A.; Manning L.; Br...	Organic milk supply in Poland: market and poli...	2019	10	Purpose: Global demand for organic milk produc...	1
4324	Zubizarreta-Gerendiain A.; Pukkala T.; Peltola H.	Effects of wood harvesting and utilisation pol...	2016	30	We studied the effects of different wood harve...	4
4325	Zuo X.X.; Lü H.Y.	Carbon sequestration within millet phytoliths ...	2011	91	Phytoliths are noncrystalline minerals that fo...	0
4326	Zuorro A.; Moreno-Sader K.A.; González-Delgado...	Evaluating the feasibility of a pilot-scale sh...	2021	6	The foreseen increase in the demand of chitin,...	1

4327 rows × 6 columns

2.3. Topic Modelling with Latent Dirichlet Allocation

The topic modelling is a tool to help the researcher discover hidden knowledge structures in document datasets (Asmussen & Møller, 2019). This allows the researcher to be informed decisions and gain insights into complex topics (Blei et al., 2003). However, choosing the right model would be challenging since the various models have different strengths and weaknesses (Jelodar et al., 2019). For example, LDA is well-known for descriptive topics learning, whereas the LSA is well-known for generating visual representations of semantic in datasets (Li & Lei, 2021; Stevens et al., 2012). After completing the pre-processing step, the TF-IDF as weighting method, which is a pre-filtering step for a statistical measure used to rate the importance of words in the content of a document set based on the occurrence of each word, then the relevant keywords in the corpus would be investigated through a bi-gram algorithm to select common phrase together with the TF-IDF algorithm to extract keywords from the abstract (Grun and Hornik, 2011).

2.4. Data Visualization

The data visualization was divined based on the group categorization in each topic from the LDA in the graph and diagram to illustrate the datasets and relationships in each dimension (Diego et al., 2020).

3. RESULTS

3.1. Data Collection Result

From collecting and selecting the document lists regarding organic agriculture and entrepreneurs in the Scopus database, the bibliographic information of the articles in each context was recorded, as well as the name of the author, title, year of publication, name of the journal, and issue. The year, number of citations, and abstract then downloaded as CSV file, which can be opened in Excel to check and select duplicate data or data without an author. There were no published years. It was found that there were 4,327 articles from 1946 to 2023. After that, files were prepared for import and analysis, including the name of the author, title, year of publication, and abstract, and added context for each article with the CSV file format, which can be imported into topic analysis programs.

Figure 4: Document computation and appeared words

```
my_nlp_experiment = nlp.setup(data=dataset, target='Abstract')
```

Description	Value
session_id	7239
Documents	4327
Vocab Size	19939

Figure 1. presents the analysis of research article publication frequency from 1946 to 2023, revealing an exponential growth in scholarly output. Initially, publication rates were low and stable, followed by gradual increases from 1980 to 2000, a significant acceleration from 2000 to 2015, and a recent surge from 2015 to 2023, reaching nearly 400 publications annually. This trend mirrors broader academic growth factors, such as increased funding, technological advancements, a more significant global research community, and more publication platforms.

Figure 2. presents the analysis of article distribution across six academic disciplines, showing notable disparities in publication volume, with Social Sciences leading at around 2,000 articles, while Multidisciplinary fields record the fewest. Secondary clusters include Business, Management, and Accounting, with 800 articles, and Economics, Econometrics, and Finance, with 700, indicating moderate output. This asymmetric distribution, with a nearly 3:1 ratio between the highest and lowest categories, may stem from varying research priorities, funding, and publication practices across disciplines, underscoring a significant concentration in Social Sciences.

3.2. Data Preparation and Data Cleaning Result

3.2.1. Indicate data categorization for analysis

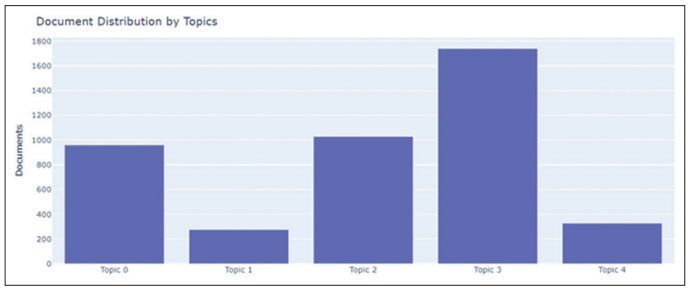
In this study, the researcher has chosen to analyze the content of the abstract section because the abstract contains an overall summary and contains important information for each article collected by preparing the data and cleaning the data, importing the.CSV file from which the data was collected into the Python

Figure 5: The results from LDA model analysis

	Authors	Title	Year	Cited by	Abstract	Sbj. A.	Topic_0	Topic_1	Topic_2	Topic_3	Topic_4	Dominant_Topic	Perc_Dominant_Topic
0	A.V V.B.; Baresel J.P.; Weedon O.; Finckh M.R.	Effects of ten years organic and conventional ...	2019	13	early vigour trait wheat composite cross popul...	0	0.708468	0.119149	0.095792	0.012793	0.063799	Topic 0	0.71
1	Aaijaz N.; Bin Ibrahim M.D.; Bin Ahmed G.	Green consumers: A growing market for SME'S an...	2010	1	study attempt gain knowledge consumer organic ...	3	0.011254	0.085088	0.183898	0.718953	0.000807	Topic 3	0.72
2	Aarnink A.J.A.; Hol J.M.G.; Beurskens A.G.C.	Ammonia emission and nutrient load in outdoor ...	2006	16	ammonia emission load outdoor run lay hen meas...	4	0.546840	0.379436	0.043079	0.028813	0.001831	Topic 0	0.55
3	Aarts H.F.M.; Conijn J.G.; Corré W.J.	Nitrogen fluxes in the plant component of the ...	2001	14	sandy area mainly use intensive dairy farming ...	4	0.558460	0.342258	0.014455	0.083324	0.001503	Topic 0	0.56
4	Abbas A.; Sajid M.B.; Sajid J.; Ahmed N.	Forecasting environmental and social benefits ...	2023	1	embody carbon new building effectively reduce ...	4	0.115941	0.587271	0.112740	0.131617	0.052432	Topic 1	0.59
...	...	...	...	...	...	...	...	...	...	...	...	...	...
4322	Zrakić M.; Jež Rogelj M.; Grgić I.	Organic agricultural production on family farm...	2017	7	croatian organic product market develop econom...	4	0.008134	0.032815	0.121155	0.835245	0.002650	Topic 3	0.84
4323	Zuba-Ciszewska M.; Kowalska A.; Manning L.; Br...	Organic milk supply in Poland: market and poli...	2019	10	purpose global demand organic milk product giv...	1	0.003496	0.172735	0.130970	0.648461	0.044338	Topic 3	0.65
4324	Zubizarreta-Gerendiain A.; Pukkala T.; Peltola H.	Effects of wood harvesting and utilisation pol...	2016	30	study effect different wood harvesting utilisa...	4	0.623315	0.245907	0.006514	0.114429	0.009834	Topic 0	0.62
4325	Zuo X.X.; Lü H.Y.	Carbon sequestration within millet phytoliths ...	2011	91	phytolith noncrystalline mineral form inside c...	0	0.806932	0.013072	0.109274	0.016638	0.054084	Topic 0	0.81
4326	Zuorro A.; Moreno-Sader K.A.; González-Delgado...	Evaluating the feasibility of a pilot-scale sh...	2021	6	foresee increase demand chitin reveal business...	1	0.060316	0.874597	0.027210	0.018608	0.019268	Topic 1	0.87

4327 rows × 13 columns

Figure 6: The topic distribution



program (Figure 3) and examining the data in the abstract section. The results indicated that there were 7,239 session data from 4,327 documents with 19,939 words (Figure 4).

3.2.2. Data processing

1. Firstly, cleaning data by eliminating unnecessary data, such as numbers, punctuation marks, symbol marks, and unnecessary words space between words, for a suitable format analysis
2. After that, the abstract has been prepared and imported into the model to analyze topics called topic modelling. Then, employed Latent Dirichlet Allocation (LDA) to define the topics. Finally, the obtained results from the LDA model show the measurement values for evaluating topics in Abstract Topic 0 – Topic 4, which is the most valuable topic in the abstract. This was considered the main topic mentioned in the abstract (Figure 5)
3. In the topic distribution, the figure shows that topic 3 was the most common topic (Figure 6). Each topic was represented by a group of keywords of the topic in the visualized text in WordCloud format, which is a visual representation of text that highlights keywords in the given content. The Wordcloud, in this study, was created from the WordCloud library in Python to present the main research content and each topic (Figure 7) (Cho et al., 2018).

3.3. Topic Modeling Results

Topic modelling is a tool for exploratory analysis of large volumes of documents (Koltsova and Koltcov, 2013; Elgesem et al., 2015), showing the overall interpretability of topics and used to assess topic quality. The Coherence Metrics technique was used to calculate statistical values and probabilities extracted from a reference library, especially focusing on the context of words to score the coherence of topics. From the probability distribution, the model can determine which topics are in a given document and which words are in the given topic by considering the distribution of words on various topics and the distribution of topics in documents. When considering the consistency of topics across topics from 0 to 6, topic 5 (Figure 8), the optimal number of topics to use for this model to maximize topic coherence. This corresponds to the 5 topics specified in the pre-processing evaluation.

3.3.1. The topic naming and topic detail

The researcher has employed the bigram's properties to predict the most common words in each topic, which the bigram feature performs better for prediction than the unigram and trigram features with overlapping structural relationships in the bibliometric analysis study on the dataset. Among the 100 most common word clusters, five topics were divided according to topic from Topic 0-Topic 4 (Figures 9-13).

Figure 7: WordCloud format from library

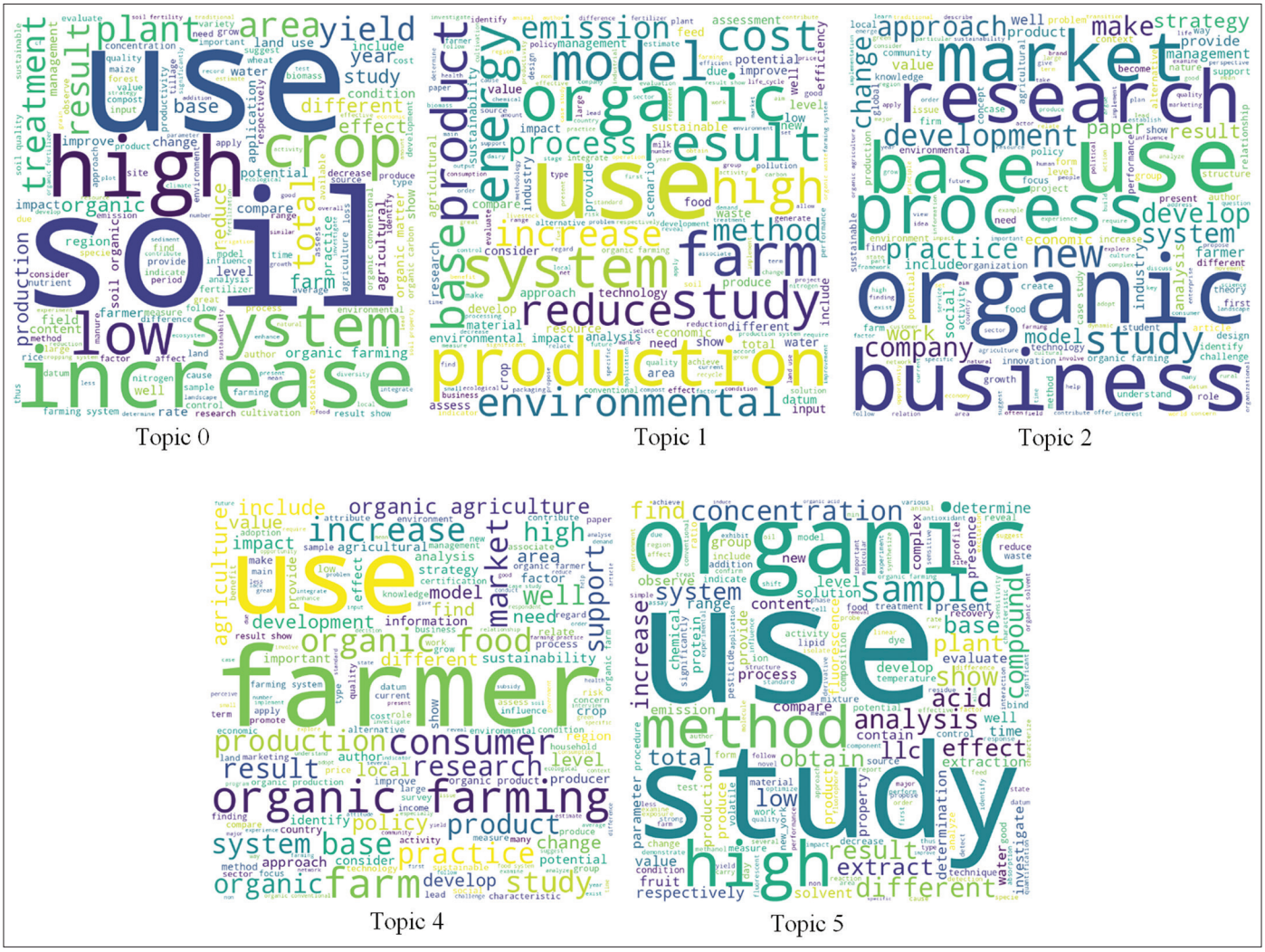


Figure 8: The appropriate topic computation results

```
from gensim.models import CoherenceModel
# instantiate topic coherence model
cm = CoherenceModel(model=lda_model, corpus=bow_corpus, texts=corpus, coherence='c_v')
# get topic coherence score
coherence_lda = cm.get_coherence()
print(coherence_lda)

0.41970792415122304
```

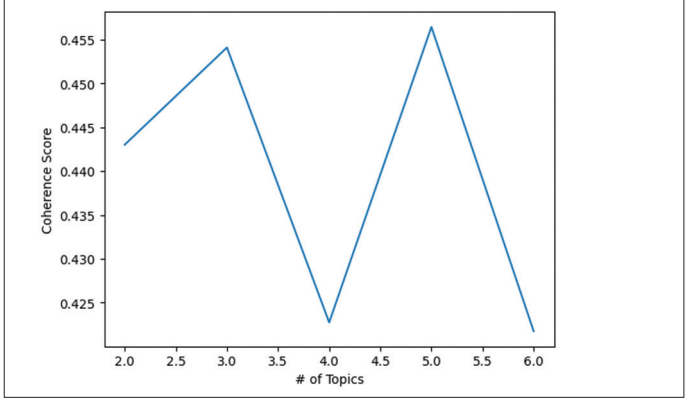


Table 1 demonstrates the topic modeling from the representative topic, suggesting that this table presents a summary or categorization of topics derived from a topic modeling analysis. In this context, "representative topic" implies that the table highlights key themes or topics that were most prominent or illustrative within a larger dataset. Each table row lists topics and details like keywords, topic relevance scores, or other metrics indicating how strongly each topic represents the dataset. This table would help readers understand the primary themes identified in the data and the relative prominence of each topic. Each topic is shown in detail below:

Topic 0, named “Soil for Organic Agriculture, “ was a concept structure that shows the relationship between soil and organic agriculture, which shows the importance of soil used in organic farming is useful in studying. This shows that studying organic agriculture requires studying soil in areas related to organic agriculture, such as land use, soil quality, and soil properties, soil improvement, adding nutrients to the soil, soil management, etc.

Topic 1, named “Environment and Organics” was a concept structure that shows the relationship between the environment and organic agriculture, which shows that the origins of organic farming are intended to preserve the environment to reduce the impact on the environment, namely agriculture using chemicals. This resulted in an excess of chemical residues in the soil mixed with water and air, continuously affecting the health of consumers

Figure 9: Word cluster from bigram algorithm in topic 0

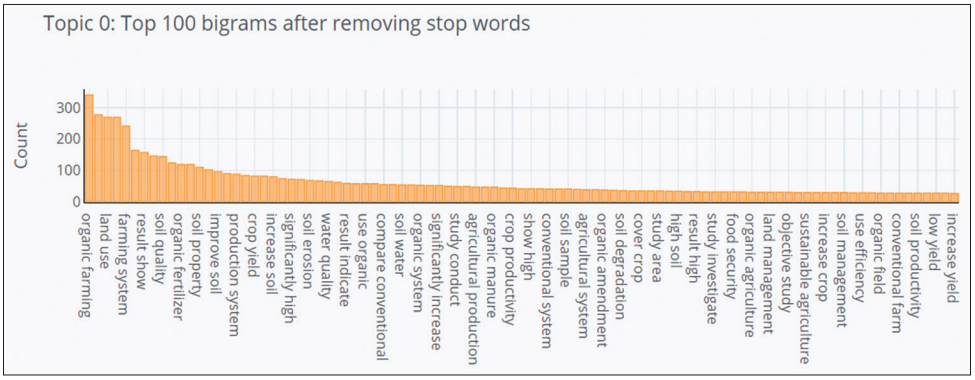


Figure 10: Word cluster from bigram algorithm in topic 1

Figure 11: Word cluster from bigram algorithm in topic 2

Figure 12: Word cluster from bigram algorithm in topic 3

Figure 13: Word cluster from bigram algorithm in topic 4

Figure 14: Topic Modeling by pyDAVis
https://dh.kku.ac.th/bibliometric/lda_organic_agriculture.html

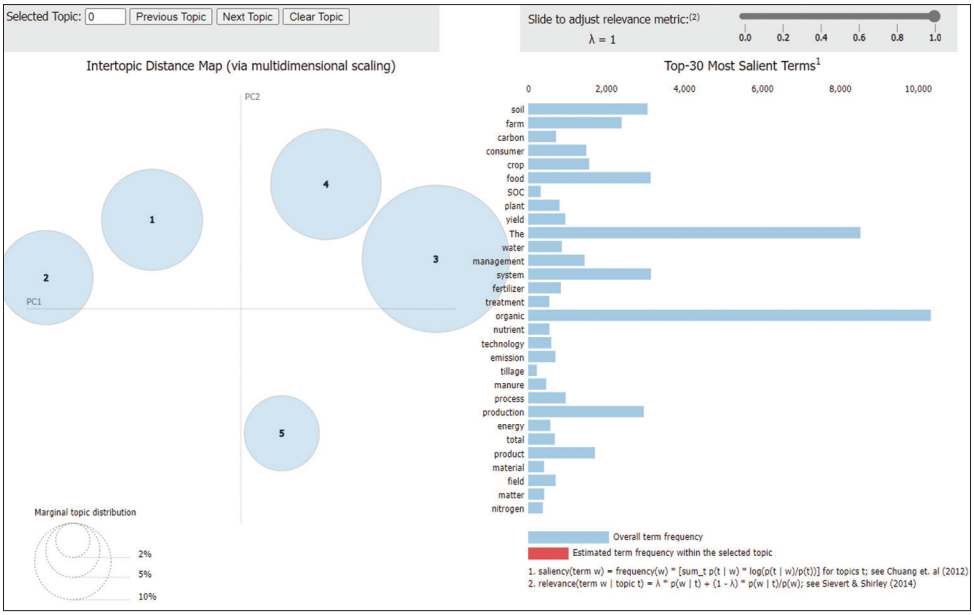
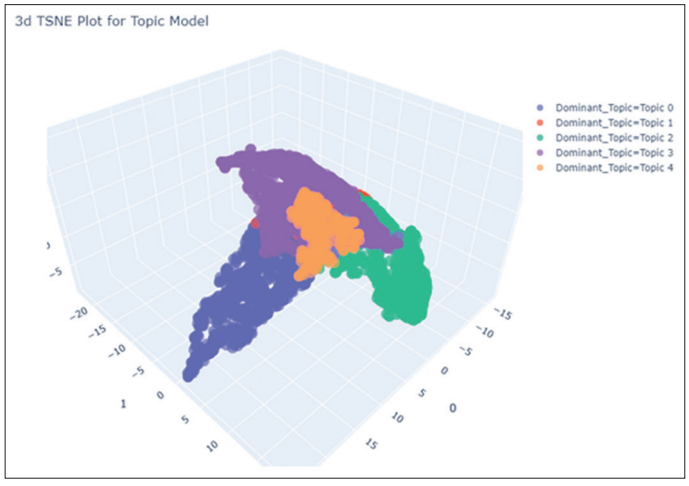


Figure 15: The Topic distribution representing in 3-dimensions



both directly and indirectly. A study of the article shows that this group of liquor information has emerged to have more knowledge and understanding of organic farming.

Topic 2, named "Organic Agriculture Business" was a concept structure that shows the relationship between business and organic agriculture to manifest how organic farming relates to the business sector. Currently, there are many forms of organic farming businesses according to the results of the topic analysis in the article, related to management, production, development, inspection, and strategy, which are business systems that classify small businesses, kitchen businesses, new business, including presentation of business results.

Topic 3 named "Organic production" and a conceptual construct representing organic production. The collected articles revealed that the most of the word groups found were in this topic cover a lot of content. Whether it is a farming system, a food system, a production system, an agricultural system, an organic system, an inspection system, organic certification, and sustainable development.

Topic 4 named "Organic use" is a concept structure that expresses the relationship between organic use and methods, which shows content related to inorganic and organic substances, chemical

residue, volatile organic, compounds organic solvents, organic acids, compound use, and concentration rate which benefits from this topic and can be used as a guideline for further studies regarding organic agriculture apart from agriculture or farming.

3.4. Data Visualization

The interactive diagramming is a highly effective way to present the results of topic models. In this regard, the pyLDAVis package (Sievert and Shirley, 2014) was used to create interactive diagrams showing the most representative topics and terms (Figure 14). The size of each circle in the diagram indicates the relevance of the topic in the corpus and topics that are close together are more similar. One of the main advantages of pyLDAvis visualization method is that users can adjust the relevance of words in a topic using a slider (Chuang et al., 2012; Sievert and Shirley, 2014). This tool offers clear and intuitive visualization and can show the relationships and strengths of each topic by displaying the words that make up each topic in a circle and horizontal bar chart.

The circles on the left panel show the overall view of the model which would allow users to easily understand the relationship between topics and their related strengths.

Meanwhile, the horizontal bar chart on the right panel shows the terms that make up each topic, which gives the user a detailed understanding of the topics.

In this study, it can be clearly seen that the five topics identified were different and belonged to different research areas. When clicking on each topic circle, the tool creates a bar graph showing the top 30 for most relevant terms for that topic. This feature allows users to quickly and concisely summarize topic relevance through the most important keywords by analyzing the words of these keywords to make it possible to categorize for all five topics.

The categories covered by these topics were highly relevant to current research topics. This can be seen from the keywords related to each topic by plotting document classifications using t-SNE (t-distributed Stochastic Neighbor Embedding). In this method, each group was represented as a probability distribution, which was basically a normal distribution to measure the distance between groups. This technique was used for 3D projections to visualize the similarity between multidimensional vectors and plot groups of similar documents (Figure 15).

4. CONCLUSION

The analysis of research on organic agriculture and entrepreneurs from the barometric data through the text analysis technique with the PyCaret library in Python, a group of words related to organic agriculture and agricultural producers were obtained. The results have expanded to reveal more knowledge, which is useful for further study in the parts that are not yet complete to be able to use it to structure knowledge more comprehensively. Especially, issues regarding the content and important concepts of organic agriculture and entrepreneurs. There were issues that farmers or entrepreneurs must study, from organic farming, production systems, development, management, and marketing to knowledge about the environment and essential organic substances to be able to do organic farming correctly according to standards and have sufficient knowledge in doing organic farming business.

In this study, the Bibliographic data were analyzed with text analysis and topic modelling techniques on abstracts from 4,327 collected articles from the Scopus database with analyzing 5 different aspects. The Python was employed to perform topic modelling and evaluate the performance of each topic. The model and present figure of the results clearly show the information which emphasizes insights and important trends from the analysis.

In the aspect of limitations in this study, this study reflected in the analysis section that the researcher had to choose stable programs and software which were able to analyze and process accurately and precisely and can be recorded and stored appropriately. In addition, the results of this study were a collection of information in the context related to organic agriculture and entrepreneurs, which was analyzed from the abstract. There might be some limit for reading research results that might not cover the entire agricultural science. Therefore, future research should explore additional fields of science to get more comprehensive results.

REFERENCES

Agricultural Research Development Agency. (2018), Modern Agriculture has Transformed into a Full-fledged Businessman. Available from: https://www.arda.or.th/knowledge_detail.php?id=24 [Last accessed on 2023 Nov 20].

Ali, M. (2020), PyCaret An Open Source Low-code Machine Learning Library. Available from: https://www.linkedin.com/pulse/pycaret-open-source-low-code-machine-learning-ali-mma-cpa-cma [Last accessed on 2023 Aug 20].

Announcement on the National Economic and Social Development Plan No. 13. (2022), Royal Gazette, 139. (Special Section 258D).

Asmussen, C.B., Møller, C. (2019), Smart literature review: A practical topic modelling approach to exploratory literature review. Journal of Big Data, 6, 93.

Blake, M., Wijetilaka, S. (2015), 5 Tips to Grow Your Startup Using SWOT Analysis. Available from: https://www.afr.com/it-pro/5-tips-to-grow-your-startup-using-swot-analysis-20150226-13pkj5 [Last accessed on 2023 Aug 20].

Blei, D.M., Ng, A.Y., Jordan, M.I. (2003), Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993-1022.

Cho, S.B., Shin, S., Kang, D.S. (2018), A study on the research trends on open innovation using topic modeling. Informatization Policy, 25(3), 52-74.

Chuang, J., Manning, C.D., Heer, J. (2012), Termite: Visualization Techniques for Assessing Textual Topic Models. In: Proceedings of the International Working Conference on Advanced Visual Interfaces. p74-77.

Diego, B.F., Mario, G., David, G., Sergio, L.M. (2020), Text mining of open-ended questions in self-assessment of university teachers: An LDA topic modeling approach. IEEE Access, 8, 35318-35330.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M. (2021), How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285-296.

Elgesem, D., Steskal, L., Diakopoulos, N. (2015), Structure and content of the discourse on climate change in the blogosphere: The big picture. Environmental Communication, 9(2), 169-188.

Grun, B., Hornik, K. (2011), Topicmodels: An R package for fitting topic

models. Journal of Statistical Software, 40(13), 1-30.
Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., Zhao, L. (2019), Latent Dirichlet Allocation (LDA) and topic modeling: Models, applications, a survey. Multimedia Tools and Applications, 78, 15169-15211.
Joo, S., Koide, S., Takeda, H., Horyu, D., Takezaki, A., Yoshida, T. (2016), Agriculture Activity Ontology: An Ontology for Core Vocabulary of Agriculture Activity. In: International Semantic Web Conference (Posters & Demos). p320-335.
Koltsova, O., Koltcov, S. (2013), Mapping the public agenda with topic modeling: The case of the Russian LiveJournal. Policy and Internet, 5(2), 207-227.
Li, X., Lei, L. (2021), A bibliometric analysis of topic modeling studies (2000-2017). Journal of Information Science, 47(2), 161-175.
National Organic Agriculture Development Committee. (2017), National Organic Agriculture Development Strategy (2017-2021). Bangkok: Office of Agricultural Economics.
Office of the Small and Medium Enterprises. (2020), Report of Strategy and Action Plan for Promoting Small and Medium Enterprises in the Agricultural Industry. Available from: https://www.sme.go.th/upload/mod_download/download-20191022060242.pdf [Last accessed on 2023 Nov 20].
Research Institute of Organic Agriculture. (2021), Report of the World of Organic Agriculture 2021. Available from: https://www.pad.moi.go.th/images/formdownload/.pdf [Last accessed on 2023 Nov 20].
Sievert, C., Shirley, K. (2014), LDAvis: A Method for Visualizing and Interpreting Topics. In: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces. p63-70.
Stevens, K., Kegelmeyer, P., Andrzejewski, D., & Buttler, D. (2012), Exploring topic coherence over many models and many topics. Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 952–961.
Wangsunthornchai, S. (1998), Using the advanced search techniques of internet search engines. Journal of Library and Information Science, 16(1), 21-36.